

EMOTION, VIRTUE, AND MORAL PERCEPTION:
A DEFENSE OF MORAL INTUITIONS

James Sias

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the
Department of Philosophy.

Chapel Hill
2013

Approved by:

Robert M. Adams

Dorit Bar-On

Geoffrey Sayre-McCord

Ram Neta

Thomas E. Hill

©2013
James Sias
ALL RIGHTS RESERVED

ABSTRACT

JAMES SIAS: Emotion, Virtue, and Moral Perception: A Defense of Moral Intuitions
(Under the direction of Robert M. Adams)

Many think that, if our moral intuitions are grounded in emotion, then they are probably not justified beliefs about objective moral values, as our ordinary practices of moral thought and discourse seem to assume. So in order to protect the epistemic status of these intuitions, some have felt it necessary to deny that our moral intuitions are grounded in emotion. But such denials are getting harder and harder to take seriously, for there is now a large, and growing, body of research in empirical moral psychology that strongly suggests that people's ordinary moral intuitions are the products of a process in which emotion figures centrally. So if we grant that this research is on point, what then can be said of the epistemic status of our moral intuitions? Skeptics insist that the empirical research forces us into the uncomfortable position of having to admit that our ordinary moral intuitions are really not what they seem—i.e., they are either not beliefs about objective moral values, or, if they are, they are not justified. In this dissertation, I resist this skepticism by constructing an account of how our moral intuitions might be trustworthy *because* they are grounded in emotion, and not *despite* this fact.

The project begins with an inquiry into the nature of emotion. On the view that I defend, emotions are complex states of mind, consisting of construals, concerns, and feelings, related to each other in a particular way. Then, after re-examining the empirical work in light of my account of emotion, I *use* the account—along with an account of how

emotion relates to moral virtue—to develop a theory of moral perception. According to the theory, virtue shapes emotions in such a way as to make them perceptions of moral value. If my argument is on point, it turns out that what is commonly assumed about the relation between emotion and the epistemic security of our moral intuitions is false—rather than threatening the epistemic security of ordinary moral intuitions, these intuitions are (or can be) epistemically secure because they are grounded in emotion.

To my children, that they might all learn to see the Good in things.

ACKNOWLEDGEMENTS

When I look back at my earliest attempts at a proposal for this dissertation, it is actually hard to believe that the project even made it this far. I had some idea of *what* I wanted to say, but no idea *how* best to say it. So first and foremost, I want to thank my advisor Robert Adams for helping me find the words, and the voice with which to say them. More important than the dissertation itself, I think this process brought me much closer to being the sort of philosopher that I want to be. I owe Bob a great measure of gratitude not only for taking me on as a student, but also for serving as a model thinker and teacher for the past three years. It was such an honor to work with him.

Dorit Bar-On might as well have been my co-advisor, as I've spent more time in her company while writing this dissertation than anyone outside of my family. She was the only member of my core committee who is not an ethicist, which turned out to be hugely important to the development and clear expression of my ideas. And if not for her unrelenting encouragement, I really doubt that I could have finished this PhD program. I cannot thank her enough for how helpful she's been to me, in so many ways.

It was a paper by Geoff Sayre-McCord that first piqued my interest in metaethics; so I give him some credit for the fact that I'm even here, doing what I'm doing. It was also Geoff's idea that I work with Bob; so in yet another indirect way, Geoff played a huge role in my becoming the philosopher that I am today. More directly, though, he is one of the most careful readers I've ever known, and my work has always benefitted from his critical eye and insightful commentary, from my first semester at UNC to the final

stages of this dissertation.

My thanks also go to the many others whose feedback helped get this dissertation and me to where we are today: to Ram Neta and Thomas Hill for being thoughtful and supportive readers; to Simon Blackburn, Doug MacLean, Laurie Paul, Jerry Postema, Ryan Preston-Roedder, David Reeve, John Roberts, Susan Wolf, and several of my fellow graduate students, for comments and questions on earlier drafts of chapters; and to Josh Knobe, Jesse Prinz, Robert Roberts, and Walter Sinnott-Armstrong, for selflessly offering their help from a distance.

A special thanks goes out to Sean Martin for reminding me, by embodying, what Hume wrote in *Treatise* 2.2.5.15. It's been a pleasure to share this with him.

Finally, I want to thank Molly. I've now completed eight total years of graduate school, during which I've written and defended two Masters theses, a proposal, and this dissertation. And this is not to mention the countless other papers, exams, and teaching responsibilities. None of this would have been possible without the constant support and inspiration that Molly has provided, and continues to provide. Our friends and family congratulate the two of us, calling this dissertation, and my Ph.D., a "team effort." But in all honesty, they have no idea just *how much* of a team effort it has truly been. I do not understand how other people manage to make it through graduate school at the distinct disadvantage of *not* having Molly in their lives.

TABLE OF CONTENTS

I.	Introduction.....	1
	1. Moral intuitions and moral skepticism.....	7
	2. The project.....	14
	3. ‘Moral intuition’ and ‘moral judgment’: clarifying two key terms.....	18
II.	What is an Emotion?.....	22
	1. Emotions as feelings.....	25
	2. Emotions as embodied appraisals.....	29
	3. Emotions as judgments.....	37
	4. My account of emotion.....	43
	5. Why this has been important.....	55
III.	Psychological Sentimentalism.....	60
	1. What ‘figures centrally’ means.....	62
	2. Moral dumbfounding.....	66
	3. Evidence of the effects of emotional stimulation upon moral judgment.....	72
	4. Evidence of the effect of emotional deficit upon moral judgment.....	80
	5. Dual-process models of moral judgment.....	89
	6. Is psychological sentimentalism an epistemic threat?.....	100
IV.	Moral Perception.....	103
	1. Why do good feelings happen to good people?.....	107
	2. Being good and feeling (the) good.....	115

3. Roberts on emotion as an epistemic aid to moral judgment.....	125
4. Reliabilism and moral reliabilism.....	130
5. Three lingering worries.....	135
6. Concluding remarks.....	140
V. Emotion and Moral Skepticism.....	142
1. Greene, Street, and the evolution of emotions.....	143
2. Nichols and the arationality of emotions.....	156
3. Sinnott-Armstrong and the clouding effect of emotions.....	161
4. Sinnott-Armstrong on ‘moral illusion’	165
5. Conclusion.....	171
References.....	175

I

INTRODUCTION

Moral intuitions are important things, to philosophers and non-philosophers alike. Outside of philosophy, moral intuitions play an integral role in our effort to make sense of ourselves, and of the world around us. We regularly make plans and decisions, form opinions about people and their actions, and behave in ways that are shaped by how things *seem* to us morally—often despite our spending little, if any, time reflecting critically upon these moral ‘seemings’.

Within philosophy, there are many who think that moral intuitions—for better or worse—are the beginning and end of moral theory. The conventional wisdom seems to be that moral theorizing begins with a relatively unstructured set of pre-theoretical intuitions about cases. From here, the theorist attempts to construct general moral principles that are supposed to cohere with those intuitions. And then those principles are subjected to a variety of tests, many of which involve weighing the principles against *more* moral intuitions about previously unconsidered cases. If ETHICAL THEORY A implies that acts of type *x* are always morally permissible, but one can imagine a case in which, *intuitively*, some token of an *x*-type action is morally wrong, this counts as a strike against ETHICAL THEORY A, and perhaps even a reason for rejecting it. Theorizing in moral philosophy *at all levels*—from some of the most abstract inquiries in metaethics to the most practical points of applied ethics—is generally thought to owe much to our

moral intuitions, and not the other way around (at least, not *as much*). Kant, for instance, adds some credence to his own moral theory by comparing its implications to his readers' intuitions about an honest shopkeeper. Mill attempts the same by comparing the implications of utilitarianism to our intuitions about such things as the desirability of happiness and the dissatisfaction of Socrates versus the satisfaction of a pig. And of course, anyone who has taken an introductory ethics course has heard these two theories criticized for failing to accord with our intuitions about, say, the permissibility of lying to a Nazi soldier or the wrongness of causing non-human animals to suffer in order to entertain humans (e.g., bull fighting). For as long as people have been doing ethics, ethical theories have often risen or fallen on the basis of their coherence with our pre-theoretical moral intuitions.

Now, to be sure, it is no small controversy whether or not this is how ethics *ought* to be done. In his influential paper "Famine, Affluence, and Morality," for instance, Peter Singer raises some serious questions about the sort of authority that we typically grant to our ordinary ways of thinking about moral issues. Of course, while defending his claim that we are morally obligated to give to those in need until the cost of giving exceeds its benefit (in terms of moral significance),¹ Singer himself appeals to his readers' moral intuitions on more than one occasion. But he also does not shy away from pointing out potential conflicts between his argument and those intuitions. And on those points, Singer is unapologetic in urging people to disregard their ordinary ways of

¹ This is the stronger version of Singer's claim. According to the weaker version, we are obligated to give until doing so involves the sacrifice of something with *any* moral significance. The difference here is unimportant, since both versions have counterintuitive implications.

thinking about charitable giving and our duties to help others, and to follow the arguments wherever the arguments happen to lead.

People do not ordinarily judge in the way I have suggested they should. [...] But given that I did not set out to present a morally neutral description of the way people make moral judgments, the way people do in fact judge has nothing to do with the validity of my conclusion. My conclusion follows from the principle which I advanced earlier, and unless that principle is rejected, or the arguments shown to be unsound, I think the conclusion must stand, however strange it appears.²

In other words, our intuitions may be at odds with Singer's argument, but as long as the argument is sound, why should our intuitions matter? Besides, as Singer urges, on a number of points, our intuitions appear to be a jumbled mess. For instance, people have the intuition that it would be wrong to refrain from saving a drowning child at the cost of ruining a new outfit, but they do not typically think the same about refraining to make an equivalent (to the cost of the outfit) donation to an organization that would use that money to save lives. Singer thinks discrepancies like this one are unjustifiable, and that they tell strongly against granting any real authority to our ordinary moral intuitions when doing ethics.

If that weren't enough to give someone pause, Singer takes his attack upon our moral intuitions even further by offering a kind of *debunking* explanation—i.e., an explanation of the origin or cause of these intuitions according to which they are unlikely to be true. The reason that we do not typically judge it wrong to refrain from giving to those in need, he explains, is that our moral intuitions depend strongly upon the strength of our social bonds with others; so the needs of people sufficiently removed from our closest social circles (family, community, subcultures, etc.) are likely to be of little

² Singer 1972: 236.

significance in our moral judgment of things. But, Singer continues, this is clearly not the way it ought to be, since morality is not so partial: “The moral point of view requires us to look beyond the interests of our own society.”³ So it’s not just that our intuitions are sometimes jumbled; more concerning is the possibility that they are systematically at odds with “the moral point of view.” (Elsewhere, Singer goes into greater detail arguing that the socio-biological origins of our moral intuitions suggest that these beliefs could be little more than untrustworthy bits of moral bias.⁴) And again, the implication is supposed to be that our ordinary moral intuitions ought to be ignored, repressed, and abandoned—by both philosophers and non-philosophers alike.

As Singer no doubt realizes, however, this is profoundly difficult. Underlying the difficulty, I think, is not just *the fact that* it is so common for us to grant our moral intuitions a great deal of authority, but more so *the reason why* we do this. And this has to do with what it is that we take moral intuitions to *be*. For one thing, at least on its face, our behavior suggests that we regard moral intuitions as genuine *beliefs*, rather than, say, expressions of attitudes or preferences.⁵ Though my wife thinks differently than I do about the superiority of cookies ‘n’ cream to all other flavors of ice cream, I do not hold her thoughts here to the sorts of norms to which we typically hold beliefs. I do not judge her to be mistaken; I do not assume that she can be convinced otherwise on the basis of things like evidence or arguments; I do not wonder if her perceptual faculties are malfunctioning; etc. On the other hand, in many cases, if she and I were to witness the

³ Singer 1972: 237.

⁴ See Singer 1981.

⁵ I address the meaning of ‘intuition’, and other terms relevant to my project, in §3 below.

same action, but from opposing moral intuitions in response, I *would* think these things about her moral intuition. So regardless of whether or not we're right to do this, it is hard to deny that we do ordinarily regard moral intuitions as beliefs.

Furthermore—still reading off of our behavior—we apparently think of our moral intuitions as beliefs about objective matters of fact. J. L. Mackie famously argued that there are no objective moral values; but before doing so, he was careful to point out that his argument is interesting precisely because our ordinary practices of judging things morally seem to take for granted that there *are* such values.

The ordinary user of moral language means to say something about whatever it is that he characterizes morally, for example a possible action, as it is in itself, or would be if it were realized, and not about, or even simply expressive of, his, or anyone else's, attitude or relation to it. ... I conclude, then, that ordinary moral judgments include a claim to objectivity, an assumption that there are objective values in just the sense in which I am concerned to deny this.⁶

Here again, it is instructive to note the differences between disagreements over things like ice cream flavors and disagreements over things like, say, the year in which the Treaty of Paris was signed. Our behavior during disagreements of the latter sort suggests that we simply take it for granted that there is a fact of the matter to which our beliefs must correspond if they are to be *true*. Our behavior during disagreements of the former sort, however, is importantly different. We might express astonishment at another's preferences, but not as if those preferences are in any sense *false*. More often than not, moral disagreements more closely resemble disagreements over things like the date of the Treaty of Paris than they do disagreements over ice cream flavors. If my wife were to believe differently than I do about the moral permissibility of abortion, we would not

⁶ Mackie 1977: 33, 35.

simply express surprise at each other's different moral taste; rather, we'd each think that the other has a *false* belief—and false in the same way that non-moral beliefs can be false. So moral intuitions are (sometimes, at least) beliefs about objective moral values, whether there *are* any such values or not.

Finally, given the sort of authority that we grant to moral intuitions, both in our ordinary lives and in ethics, we apparently assume that our moral intuitions are sometimes *justified* as well.⁷ After all, it would be odd for us to lean upon these beliefs as heavily as we so often do if we did not think that they were had for good reason(s). It would be odd, that is, for us to so regularly make plans, form opinions, and behave in ways based upon beliefs with the epistemic status of mere guesses or worse. And it would be remarkable if it turned out that the whole history of Western moral philosophy was riddled with philosophers resting their cases (even slightly) upon unfounded moral hunches. Ordinarily, whenever we are skeptical of the justificatory status of a belief, whether our own or someone else's, we do exactly as Singer urges—we either discard it, or at least subject it to some degree of suspicion until its epistemic status is resolved. With our moral intuitions, however, we generally only do this if we are presented with some compelling reason to doubt their justification—which suggests that our default is to regard them as justified.

In short, though we may sometimes deny this when pressed, we nonetheless *go on as if* our ordinary moral intuitions are instances of *moral knowledge*. This explains the profound difficulty of disregarding them, as Singer urges us to do. In fact, as I'll go on to

⁷ I do not have any particular theory of epistemic justification in mind here. I mean only to suggest that we often behave as if our moral intuitions have a sufficient degree of *whatever it is* that separates knowledge from mere true belief.

argue in this dissertation, moral intuitions resemble instances of *perceptual* beliefs. We *trust* the process by which our ordinary moral intuitions are formed like we trust the process by which our ordinary perceptual beliefs are formed. It is no wonder, then, that Singer's recommendation that we occasionally disregard our intuitions can seem so foreign and upsetting. No matter how compellingly I've *argued* that *p*, if you *see* things in a way that implies *not-p* (and you don't have any additional reasons for doubting the veracity of your perceptual experiences), you'll likely go on as if it is the argument, and not the perception, that must be disregarded.

1. Moral intuitions and moral skepticism

While we apparently treat our moral intuitions like they are bits of moral knowledge, there are some who deny that such knowledge is possible (even if we grant that there are moral values), and their reasons for doing so have directly to do with epistemic status of moral intuitions. According to Walter Sinnott-Armstrong, the "deepest challenge in moral epistemology" is essentially an application of the famous regress argument:

Someone is justified in believing something only if the believer has a reason that is expressible in an inference with premises that the believer is already justified in believing. This requires a chain of inferences that must continue infinitely, close into a circle, or stop arbitrarily. Academic skeptics reject all three options and conclude that there is no way for anyone to be justified in believing anything. The same regress arises for moral beliefs.⁸

As Sinnott-Armstrong goes on to acknowledge, the simplest way to avoid the regress is to argue that there is some belief in the chain of inferences that is justified *non-inferentially*, i.e., justified without needing to be inferred from some other justified belief. What would this look like in the case of moral beliefs? Well, presumably, the non-

⁸ Sinnott-Armstrong 2006: 340.

inferentially justified belief would have to be either *another* moral belief or a *non*-moral belief. But many think that it cannot be a non-moral belief, because non-moral beliefs cannot be taken to justify moral beliefs, else we risk violating Hume's famous Is-Ought prohibition.⁹ So if moral knowledge is possible, it will have to turn out there are some moral beliefs that are non-inferentially justified.

Now, the good news is that there is an entire tradition in moral epistemology, known as *ethical intuitionism*, according to which some moral beliefs are non-inferentially justified.¹⁰ As we can see, then, ethical intuitionism offers an immediate reply to the regress argument: if some moral beliefs are able to be justified without receiving any kind of inferential support from other beliefs, then there is a legitimate place for the regress to terminate, and so a legitimate ground for moral knowledge claims. Furthermore, if it turns out that *enough* of our ordinary moral intuitions are among those that are non-inferentially justified for the process by which those intuitions are formed to count as reliable, then perhaps we're *right* to go on as if our moral intuitions are instances of moral knowledge, as it appears we do.

The bad news, however, is that ethical intuitionism is thought by many to be a hopelessly flawed view.¹¹ The main point of contention seems to be the claim made by

⁹ For more on this, see Shafer-Landau 2004: 118-121.

¹⁰ Ethical intuitionism is often understood as a view that takes a position upon not only the epistemology of moral beliefs, but the metaphysics and semantics of morality as well. Typically, intuitionists are realists and non-naturalists about moral values, and they are cognitivists and descriptivists about moral claims. On other occasions, however, intuitionism is understood only as the moral epistemological view that I describe in the text.

¹¹ Much of the criticism that intuitionism has received over the years has focused upon its close association with ethical non-naturalism (see n. 10). Many (including most intuitionists) see the view as committed to the belief that moral values are things metaphysically distinct from the natural world. And

many intuitionists that moral beliefs are justified non-inferentially when their contents—moral propositions—are *self-evident*, i.e., knowable solely on the basis of an understanding of the terms and concepts of which they are made up. This is how many believe that basic truths in *mathematics* are known, for instance, and so it comes as no surprise that intuitionists have for centuries likened moral knowledge to knowledge in mathematics. But there are a number of problems with this approach. For one thing, depending upon how high one sets the standards for *understanding* the relevant terms and concepts, it may turn out that far fewer people have moral knowledge than we might like to admit. And this would be ironic indeed, given that intuitionists often promote their view as amounting to a defense of moral knowledge as “common sense.”¹²

Another problem with the analogy that intuitionists sometimes try to draw between moral and mathematical knowledge is the fact that there seems to be so much more disagreement over moral matters than there is over matters of mathematics. Now, to be sure, it would be a mistake to think that it simply *follows* from the fact of widespread moral disagreement that there cannot be self-evident moral propositions. But a slightly weaker charge may nonetheless have a bit more force: even if self-evidence does not *guarantee* agreement of beliefs, it surely lends itself to such agreement, as we see in mathematics; so if there is significantly more disagreement in ethics than there is in mathematics, this would be a *prima facie* reason to doubt that moral propositions admit of self-evidence; and in fact, there *is* significantly more disagreement in ethics.

this is controversial, to say the very least. But since my interests here are epistemological, I'll ignore whatever metaphysical controversies there may be associated with ethical intuitionism.

¹² See, e.g., Stratton-Lake 2002: 1.

Finally, a perhaps more interesting difference between math and morality is the oft-noticed fact that beliefs about the latter seem to engage us motivationally in ways that beliefs about the former do not. Whatever the relation between *having a moral belief* and *being motivated to act accordingly* is supposed to be exactly, the fact that there is some close relation has led many to think that the process by which moral beliefs are formed must be fundamentally different than that by which mathematical beliefs are formed. It cannot be the purely rational process that many intuitionists describe. What, then, is the unique process by which beliefs about moral values are formed, some of which are supposed to be justified? For reasons I've already noted, intuitionists should no longer be comfortable falling back upon the analogy to knowledge in mathematics. But then it is unclear what intuitionists *can* say. As Wedgwood notes,

Many ... say little more than that we have some cognitive faculty—sometimes called “intuition” or “reason” or “conscience”—which enables us to come to know and have justified beliefs in normative propositions; but they rarely give any account of how exactly this alleged faculty operates, or how it could serve as a reliable source of knowledge, or what could justify us in relying on it. For this reason, anti-realists often accuse realists of failing to meet certain crucial demands for explanation.¹³

And because of their unfortunate silence on this point, Wedgwood goes on to note, intuitionists open themselves up to Mackie's famous attack on grounds of queerness:

If there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe. Correspondingly, if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else.¹⁴

Intuitionists *would* have a handy response to the regress argument, if only they could explain how it is that moral beliefs could be non-inferentially justified. But the

¹³ Wedgwood 2007: 225.

¹⁴ Mackie 1977: 38.

way in which they have typically done this—via an analogy to the self-evidence of propositions about other subjects, like mathematics—seems to many to be devastatingly problematic, ultimately *raising* more questions than it answers. Despite a recent swelling of interest in intuitionism, I think there are many who would still echo G. J. Warnock’s 1967 verdict: “Intuitionism seems, in retrospect, so strange a phenomenon—a body of writing so acute and at the same time so totally unilluminating—that one may wonder how to explain it, what its genesis was.”¹⁵

So things are not looking particularly good for the status of our ordinary moral intuitions. As I said above, we *go on as if* they are justified beliefs about objective moral values. But in light of the challenge posed by the regress argument, and the apparent difficulty of accounting for how moral propositions could be self-evident, it just is not clear how moral intuitions could be justified. Given the way in which they are formed, it does not seem as if our moral intuitions are ordinarily the conclusions of any kind of *inference* from justified beliefs. And the prospects of *non-inferential* justification of moral intuitions now seem pretty bleak.

Making matters worse, apparently, is the fact that recent work in empirical moral psychology strongly suggests that, far from being the products of some sort of *rational faculty*, our ordinary moral intuitions are in fact the effects of *emotion*. So ethical intuitionism—or at least, many versions of the view—may suffer from a kind of empirical inadequacy as well. Summarizing the past 20 years or so of research into the

¹⁵ Warnock 1967: 16. For what it’s worth, Warnock’s own explanation of the genesis of intuitionism was that intuitionists were lacking in both curiosity and doubt.

neuro-psychological underpinnings of moral judgment, psychologists Joshua Greene and Jonathan Haidt write,

[H]ow do we decide that someone else has done something wrong? [...] Recently, [...] findings from several areas of cognitive neuroscience have begun to converge on an answer: emotions and reasoning both matter, but automatic emotional processes tend to dominate.¹⁶

Thorough discussion of this research will have to be put off until chapter III. For now, I want only to explain *what* the research seems to imply, and *why* this matters to the foregoing discussion of the epistemology of moral intuitions. Roughly, the research implies that our ordinary moral intuitions—the ones we form while watching the news, or talking with a friend, or even reading a piece of fiction—are the products of a process in which emotion figures centrally as a cause. In other words, at least in ordinary cases, *we have the intuitions that we have because we respond emotionally to things in the ways that we do.*

This is significant to the epistemic status of our moral intuitions because there is a long-standing assumption, shared by philosophers and non-philosophers, according to which emotions are at best *distractions* or *obstacles* to clear thinking. As Alison Jaggar notes, “Within the Western philosophical tradition, emotions have usually been considered potentially or actually subversive of knowledge.”¹⁷ In a 2007 article entitled “Should We Trust Our Moral Intuitions?” Singer describes some of the relevant empirical psychological research—agreeing that it shows our moral intuitions to be grounded in emotion—and concludes, “[T]hese findings should make us more skeptical about relying

¹⁶ Greene and Haidt 2002: 517.

¹⁷ Jaggar 2008: 378.

on our intuitions.”¹⁸ And of course, outside of philosophy, we regularly remark about the distorting influences of emotions upon our thoughts and behavior. We say things like, “I was *blinded* by anger,” “He was *overcome* with fear,” and the like. When the stakes of a particular belief or decision are relatively high, we are careful to recommend that people “keep a cool head,” and try not to let their emotions “get the best of them.” So even if moral beliefs *could* be non-inferentially justified, it would be easy to conclude on the basis of this empirical research that our ordinary moral intuitions are *not*.

Let me pause here to briefly take some stock, before going on in the next section to describe my project. On the one hand, it is hard to deny that moral intuitions do, in fact, play the sorts of roles described above, both in the ordering and conducting of our everyday lives and in the task of doing ethics. But on the other hand, I think it would be equally hard to deny, as Singer and others have pointed out, that our moral intuitions sometimes exhibit signs of being *undeserving* of the sort of trust that we apparently give to them. Cutting even deeper than this is a kind of origins problem for moral intuitions: there is a large, and growing, body of empirical research into the psychological origins of our ordinary moral intuitions, and the evidence seems strongly to suggest that these intuitions are the products of a process in which emotion figures most prominently as a cause. And when this is coupled with the popular assumption that emotion is at best a distraction or obstacle to clear thinking, it looks as if we’ve got a serious problem on our hands. It is not just that our moral intuitions are sometimes untrustworthy—this could probably be said of all kinds of beliefs, under certain conditions. Rather, if the empirical research is on point, it appears as if moral intuitions are *systematically* untrustworthy.

¹⁸ Singer 2007.

Something has got to give. If the empirical research does indeed establish what it purports to establish—and in chapter III, I argue that it does—then we must either (i) somehow abandon the trust we apparently place in our ordinary moral intuitions, and also cease using them even as data points for moral theory; or else (ii) reject the popular assumption that our moral intuitions cannot be trusted if they are grounded in emotion.

2. The project

This dissertation is an attempt to defend the epistemic status of our moral intuitions in light of the recent empirical evidence of their being grounded in emotion. At the end of the day, there may be *other* reasons for being “skeptical about relying on our intuitions,” as Singer puts it; but the fact that they are ordinarily the products of a process in which emotion figures centrally, I’ll argue, is not one of them. And this is because the popular assumption is false: it is not the case that emotion could only be a distraction or obstacle to the formation of justified beliefs about objective moral values. On some occasions, I’ll argue, a person’s intuitions will count as justified beliefs about objective moral values *because* they are grounded causally in emotion.

On my account, the sort of justification that ordinary moral intuitions can enjoy is indeed non-inferential—and so, if my account is successful, there may be hope for intuitionists to stop the regress after all. In defense of the non-inferential justification of (some) moral intuitions, I will follow-up on an analogy at which I hinted earlier, i.e., an analogy between moral intuitions and ordinary perceptual beliefs. Ordinary perceptual beliefs count as non-inferentially justified, but not because their contents are *self-evident* in the sense described above. Rather, they count as non-inferentially justified because of

the process by which they are ordinarily formed. Likewise, I'll argue, the process by which moral intuitions are ordinarily formed is one that can confer upon them a similar sort of non-inferential justification. (I have more to say about justification in moral and non-moral cases in chapter IV.) Interestingly, I think my account of the process by which moral intuitions are formed will also provide me with unique explanations for both the amount of disagreement over moral matters and the apparent connection between moral judgment and motivation.

Furthermore, as far as the “special faculty of moral perception or intuition” is concerned, I intend to posit nothing more “queer” than the emotions with which we are already quite familiar. To be sure, I will assume that there are objective moral values, something that Mackie obviously thought queer. But as is evident from the quotation above, he, too, helps himself to an assumption that objective moral values exist, if only to speculate about the relative queerness of whatever faculty allows us to be aware of them. In fact, moral skeptics routinely begin their arguments by assuming that there are objective moral values—after all, if we do not at least assume that there is an x to be known, why bother arguing that knowledge of x is impossible? And since my aims in this dissertation are purely epistemological, I intend to do the same, i.e., simply assuming that some variety of moral realism is the case. Only then can I go on to explain, over the course of chapters II-IV, how it is that ordinary moral intuitions are, or can be, justified beliefs about real, objective moral values.

Here is how the project will unfold. In the next chapter, I'll ask, “What is an emotion?” Given how many philosophers and psychologists have speculated about the

impact of emotion upon the epistemic status of moral intuitions, it is remarkable how few of them have had anything to say about the actual nature of emotion. Historically, theories of emotion have fallen into one of two camps. On the one hand, *feeling theorists* argue that emotions are essentially just that—*feelings*. These feelings are typically bodily in nature, but strictly speaking, a feeling theory of emotion could allow that there are non-bodily emotional feelings. On the other hand, *judgment theorists* claim that emotions are a species of *judgment* or belief—typically, *evaluative* judgment or belief. More recently, philosophers have grown interested in the prospects of hybrid, or compound, theories of emotion—i.e., theories according to which emotions exhibit features similar to both feelings and judgment. After briefly reviewing two prominent feeling theories and one judgment theory, I'll go on to defend a hybrid theory according to which emotions are complex mental states, consisting of what I will call construals, concerns, and feelings, all related to each other in particular ways.

Then, in chapter III, with a working account of the nature of emotion in hand, I'll review some of the empirical research mentioned above, especially those studies that I think most strongly implicate a causal relationship between emotion and moral intuitions. Some of the evidence will come from neuroimaging studies, some from social psychological studies, and some from studies in abnormal psychology. Together, however, I believe it all amounts to a strong case for what is sometimes called *psychological sentimentalism*, which I define as the view that ordinary moral intuitions

are the products of a psychological process in which emotion figures centrally as a cause.¹⁹

Chapter IV is the centerpiece of the project. In this chapter, I bring the discussions from chapters II and III together in defense of the claim that moral judgment *can be* a process that confers non-inferential justification upon moral intuitions. My argument depends heavily upon the theory of emotion that I defend in chapter II, and also upon an account of the relationship between emotion and moral virtue that I develop in chapter IV.

Finally, in chapter V, I respond to four arguments against thinking that moral intuitions are justified beliefs about objective moral values. The first three arguments directly address the significance of emotion to moral judgment. One way or another, they all claim either that moral intuitions *cannot be*, or at least that they *are probably not*, justified beliefs about objective moral values if they are causally grounded in emotion. The fourth argument, also from Sinnott-Armstrong, has to do with the susceptibility of moral judgment to what are sometimes called *framing effects*. Rather than arguing that moral judgment is *not* susceptible in this way, I'll argue that these studies actually provide an interesting sort of confirmation of my overall view. I conclude this chapter, and the dissertation, by speculating about further avenues of research that may be opened up by my view.

¹⁹ Invoking the name 'sentimentalism' in this way naturally invites the following question: What sort of relationship am I assuming between emotion and *the sentiments*? There are a few things to say here. I think it's clear that modern moral philosophers like Hume and Smith meant something different by 'sentiment' than I will mean in this dissertation by 'emotion', though I do think they would include emotions among the sentiments. Basically, this is how I am thinking of the relation between emotions and sentiments as well, but I do not think that much will turn on this issue. Here I am simply following recent convention in defining 'psychological sentimentalism' as the view that emotion figures centrally in moral judgment.

3. ‘Moral intuition’ and ‘moral judgment’: clarifying two key terms

Before proceeding, I should clarify my use of two key terms. This will be especially important for following the discussion in chapters III-V.

The term ‘moral intuition’ is sometimes used to refer to a kind of *capacity*, i.e., an instinct or insight into moral matters. Indeed, in general, the term ‘intuition’ is often used in this way, as when my mother-in-law suggests that she has a “mother’s intuition,” by which she apparently means a kind of insight into the thoughts and behavior of her children. Other times, however, the term ‘moral intuition’ is clearly used to refer to a kind of propositional attitude, like belief. For instance, we sometimes say things like, “I have the intuition that such-and-such would be wrong,” or, “I have conflicting intuitions about what would be the right thing to do.” Throughout this dissertation, as I have done here in the Introduction, I will use the term ‘moral intuition’ only in this latter sense, i.e., to refer to a *moral belief* with certain distinctive features.

Interestingly, while both philosophers and psychologists regularly discuss moral intuitions, the two groups seem to isolate different—though related—features that they take to be characteristic of moral intuitions. On the one hand, philosophers generally focus on the *non-inferential* nature of these beliefs. Saying nothing at all (for now) about the justificatory status of moral intuitions, philosophers nonetheless generally concede that moral intuitions are distinctive in that they are not formed on the basis of any kind of inference from other beliefs. As I’ve already noted, moral intuitions are typically where moral reasoning *begins*, both in our ordinary lives and in ethical theory. On the other

hand, psychologists often note the speed and effortlessness with which moral intuitions are formed. Jonathan Haidt explains,

Commentators on intuition have generally stressed the fact that a judgment, solution, or other conclusion appears suddenly and effortlessly in consciousness, without any awareness by the person of the mental processes that led to the outcome.²⁰

I think it is clear enough that philosophers and psychologists are talking about the same thing, just with different concerns in mind. Since philosophers are more likely to be concerned about the *epistemic status* of moral intuitions, it makes sense that they would focus on the fact that moral intuitions are not formed on the basis of any sort of inference. And since psychologists are more likely to be concerned about the causal mechanisms underlying moral intuitions, it makes sense for them to focus on the automaticity and effortlessness with which they are formed. But at the end of the day, both groups are talking about a set of moral beliefs distinguished by the fact that they are *not* the products

²⁰ Haidt 2001: 818. I actually think Haidt is unclear in his use of the term ‘moral intuition’. At times, he uses it to refer to the psychological process underlying the sudden appearance in consciousness of an attitude or belief: “Moral intuition is therefore the psychological process that Scottish philosophers talked about, a process akin to aesthetic judgment” (*ibid.*). At other times, it seems he is using it to refer not to the psychological process underlying this appearance, but to the appearance itself: “*moral intuition* can be defined as the sudden appearance in consciousness of a moral judgment” (*ibid.*). To make matters even less clear, while he here uses the term ‘moral judgment’ to refer to the attitude or belief that appears in one’s consciousness, elsewhere, he clearly uses the term ‘moral judgment’ to refer instead to the psychological process underlying this appearance—a process which, again, he sometimes calls ‘moral intuition’.

I say this not to pick on Haidt, but to highlight the fact that it is awfully easy to let one’s use of these terms go unchecked; and as a result, one risks making claims about one thing—e.g., the attitude or belief, the appearance of the attitude or belief in consciousness, the process underling the appearance, etc.—that would be more accurately made about something else. If, for instance, I were to say that *moral judgment is based upon emotion*, but I were also to vacillate in this way with my use of the term ‘moral judgment’, my audience is liable to interpret me as making any number of importantly different claims. I might be saying that the psychological process by which moral intuitions (as beliefs) are formed is one in which emotion figures centrally. Or I might be saying that, in certain contexts, moral judgments and emotions are identical. Or I might be saying something completely different. It is for precisely this reason that it is important for me to clarify my use of these terms early, and then to stay consistent in my use of them throughout the entire dissertation.

of a cognitively effortful (and so, slower) process of inference or conscious reasoning.²¹ I, too, will use ‘moral intuition’ to refer to moral beliefs with these features—i.e., moral beliefs formed relatively quickly and effortlessly, not on the basis of any inference.

The term ‘moral judgment’ is perhaps even more diverse in its usage among philosophers and psychologists. Like ‘moral intuition’, ‘moral judgment’ is sometimes used to refer to a belief or propositional attitude of some sort—as when we say, “It’s my judgment that it would be wrong to tell a lie.” In examples like this, ‘(moral) judgment’ is apparently interchangeable with ‘(moral) belief’. And like ‘moral intuition’, we also sometimes use ‘moral judgment’ to refer to a kind of capacity, like sight or creativity. Used in this way, moral judgment is something you exercise, something in which you engage; and some can be better at it than others—e.g., “She has excellent judgment in these (moral) matters.”

Other times, the term ‘moral judgment’ is used to refer to a kind of mental *act*—i.e., the act of forming a particular moral belief or attitude. At still other times, ‘moral judgment’ is used to refer not to a belief or the act of forming a belief, but rather to the *content* of a belief—something like a Fregean *thought*. In this sense, moral judgments are expressible in sentences, bearers of semantic value, and also truth-evaluable. And finally, some (especially psychologists) use the term ‘moral judgment’ to refer to a kind of psychological process—specifically, the term is often used to refer to the psychological process responsible for the production of moral intuitions (see, e.g., n. 20).

²¹ Note that, for all I say here, moral intuitions may yet be the process of *some kind of* inferential process—just, perhaps, one that takes the form of an unconscious (and so, cognitively effortless) analysis.

For my part, I will only use the term ‘moral judgment’ in this last way. Moral intuitions are moral beliefs that are formed quickly and effortlessly, not on the basis of any inference; and moral judgment is the psychological process that is ordinarily (causally) responsible for their production. To be clear, I do not assume that our moral intuitions are always formed in exactly the same way. But I do assume that there is *something* that can be identified as the process by which they are *ordinarily* formed. I’ll use ‘moral judgment’ to refer to *that* process, whatever it might involve.

II WHAT IS AN EMOTION?

As I explained in the last chapter, the focus of my project is upon the role of emotion in moral judgment, and the implications of this upon the epistemic status of moral intuitions. Since it will be difficult to draw any conclusions about either of these issues without first having some idea of what an emotion is, that is where my inquiry will begin: What is an emotion?

Let me start by identifying a few of what I take to be ordinary features of emotion. First of all, emotional experiences typically involve feelings of various sorts. For instance, fear is often accompanied by such things as an increased heart rate, shortness of breath, trembling limbs, and the sensation of being frozen, unable to move, etc. Second, there are conceptual ties between emotions and the conditions that typically elicit them. In paradigmatic cases, for instance, fear is elicited by dangers; anger by offenses; grief by losses; etc. And these elicitors figure into our conceptions of the corresponding emotions—e.g., danger figures into our conception of *what fear is*. Third, emotions are intentional states, in the sense that they are *directed at*, or *about*, things in the world.¹ When we are afraid, there is something we are afraid *of*; when angry, there is something we are angry *with* or *at*; and so forth. And finally, emotions are sometimes

¹ This point about the intentionality of emotion is often taken to mark the distinction between emotions and *moods*. Emotions take objects; moods do not. Fear and anxiety are both affective states, but unlike fear, anxiety can occur without an object. That is, we can be anxious *about nothing at all*—in such a case, the anxiety is a mood, not an emotion. Fear, however, must always be *of* something.

subject to evaluation—both rational and moral. It is *right* (in some sense) to fear a ferocious bear, and *wrong* (in a related sense) to fear Yogi bear. If someone is not ashamed of some terrible behavior, we regard the lack of shame as a moral failure in itself. So here are four ordinary features of emotion: emotions are accompanied by feelings; there are conceptual ties between emotions and their elicitors; emotions are intentional; and emotions are sometimes subject to evaluation. There certainly may be more features that emotions share, but for now, I'll assume that a satisfactory account of the nature of emotion should at least be able to account for these four.

So what sort of thing is an emotion? What sort of thing could have all of these features? In this chapter, I'll argue that emotions are complex states, consisting of three things—construals, concerns, and feelings—bearing particular relations to each other. As I'll explain later (in §4), a construal is a type of mental state in which one thing is seen *in terms of* something else. Construals are cases of *seeing-as*—e.g., when I see a coffee cup *as empty* (as opposed to seeing it *as dirty*, or *as a gift from my wife*), I construe the cup in terms of the concept empty. Often times, my construal of a thing, i.e., the way I see it, will either clash or accord with my concerns. My construal of the coffee cup as empty, for instance, will clash with my concern, my desire, for more coffee. My construal of a snake as a threat to my wellbeing will clash with my concern for my wellbeing. And finally, sometimes, the relation between construals and concerns will give rise to feelings of various sorts. When these three things come together like this—i.e., when construals relate to concerns in ways that give rise to feelings—the result is an emotion.

This may sound more complicated than we might have expected from an account of something as commonplace as emotion, but as I'll attempt to show in the next few sections, the account is nonetheless well motivated, as it does a better job than other emotion theories of accommodating the four features of emotion that I describe above. In §§1-3, I'll discuss three prominent theories of emotion—William James' feeling theory, Jesse Prinz's embodied appraisal theory, and Robert Solomon's judgment theory, respectively. Each theory, I'll argue, does a better job than the one before it of accounting for the four ordinary features of emotion, but none can accommodate all of them. And it will be important to see not only *that* they fail to accommodate them, but also *how* and *why* this is the case, for this will allow us to get a clearer picture of what a satisfactory account will have to look like.

Before getting into it, though, I want to issue two quick caveats. First, to those who doubt that a coherent and unifying theory of emotion is possible,² I'll have basically nothing to say. I assume that there is something that unites the emotions, and I'll explain what I think that is in §4. But I'll certainly not be running through the myriad things that have ever been called “emotions,” and considering whether or not they can be accounted for by the theory I endorse. Throughout the chapter, my discussion will focus upon only the most paradigmatic of emotions—e.g., fear, anger, grief, etc. If *anything* is an emotion, these are emotions.

Second, I am only interested in getting a sufficient handle on the nature of emotion *so that* I can go on in subsequent chapters to explore the role that emotion plays in moral judgment. So my aim here will not be to establish a complete theory of

² See, e.g., Rorty 1980 and Griffiths 2004.

emotion, i.e., one that addresses everything we might expect a theory of emotion to address. That would require much more space than a single chapter, and likely involve discussions of issues not obviously relevant to moral judgment (e.g., similarities and differences between emotions, moods, and other affective phenomena; the roles played by different parts of the brain in the experience of particular emotions; etc.).

1. Emotions as feelings

In his seminal work *The Principles of Psychology* (1890), William James defends a theory of emotion according to which emotions are perceived (i.e., felt) changes to a subject's physiology. When I am angry, for instance, my face and neck may feel warmer, and my heart may seem to beat faster and with greater force. According to a feeling theory of emotion, my perception, or *feeling*, of these physiological changes (and others) *just is* my anger.

Our natural way of thinking about these [...] emotions is that the mental perception of some fact excites the mental affection called the emotion, and that this latter state of mind gives rise to the bodily expression. My theory, on the contrary, is that *the bodily changes follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur IS the emotion.*³

In support of his theory, James has us imagine abstracting away all such feelings, and then asks: what is left of the emotion? If we take away the warming of my face and neck, the increased heart rate, and any other such “*feelings of its bodily symptoms, we find that we have nothing left behind*, no ‘mind-stuff’ out of which the [anger] can be constituted.”⁴ In fact, in the case of anger, James thinks, once we take away all feelings of associated physiological changes, all that will be left is a certain sort of judgment, i.e.,

³ James 1950: 449-450; italics in original.

⁴ James 1950: 451.

“some cold-blooded and dispassionate judicial sentence, confined to the intellectual realm, to the effect that a certain person or persons merits chastisement for their sins.”⁵

It is certainly true that emotions often manifest themselves physiologically, and that these physiological changes are a part of what it is to feel the emotion. Anger has a certain feel to it, one that often includes an increased heart rate, warming of the face and neck, and tightening of certain muscles. Feeling theorists may even be right to suggest that each emotion has its own unique physiological signature. James writes,

The various permutations and combinations of which these organic activities [i.e., the physiological changes our perception of which, he thinks, constitutes an emotion] are susceptible make it abstractly possible that no shade of emotion, however slight, should be without a bodily reverberation as unique, when taken in its totality, as is the mental mood itself.⁶

Fellow feeling theorist Carl Lange agrees, adding that it is their unique physiological signatures that allow us to distinguish between emotions. What distinguishes anger from fear, for instance, is the fact that the former involves swelling of the blood vessels, an increase in the secretion of saliva, and the onset of irregular breathing, while the latter involves heart palpitations, dryness of the mouth, and trembling of limbs.⁷

So maybe we should concede that James’ theory has indeed captured one of the four ordinary features of emotion I described earlier—i.e., that they are accompanied by feelings of various sorts.⁸ I do not, however, think that his theory can account for the other three features. For starters, while James does acknowledge that emotions are

⁵ James 1950: 452.

⁶ James 1950: 450.

⁷ Lange 1967: 46-54.

⁸ In §4, I’ll suggest that emotional feelings are sometimes more psychological than physiological.

elicited by things, he seems not to recognize the conceptual ties between emotions and their elicitors. Contrary to what James and Lange suggest, what distinguishes anger from fear is not just the set of differences between their physiological signatures, but also—and, I would argue, *primarily*—the relations they bear to such things as dangers and offenses, respectively.

Furthermore, James' feeling theory is notorious for its failure to account for the intentionality of emotion. Of course, the feelings with which James identifies emotions are intentional in a sense. They are feelings *of* something—namely, physiological changes. But this is obviously not what anyone has in mind when they claim that emotions are intentional states. When I am lied to, for instance, my anger is direct *at the liar*, and not anything going on in my body. So even if James' theory allows *some* sense in which emotions are intentional, he has obviously not accounted for the intentionality of emotion in the right way. And it's not clear how he could: how could the feeling of certain physiological changes ever be directed at *something else*?

James wrongly assumes that the object of an emotion is just whatever *caused* it, but as many have since realized, the relationship between an emotion and its object cannot be a causal one. Even if emotions are sometimes directed at their causes, this is not always the case. Suppose Susan tells me that she overheard John, a colleague of mine, slandering me, but she is lying—John did not actually slander me. I will be angry, and my anger will be directed at John. He is the object of my anger. But in what sense is John the *cause* of my anger? Maybe John doesn't even exist—I've fallen victim to an elaborate hoax, set up to get me to believe that I have a colleague named John who

slanders me behind my back. Yet I am genuinely angry, and angry *at John*. This case may be contrived, but I think it nonetheless highlights the fact that the *object-of* and *cause-of* relations operate fundamentally differently with respect to emotion. (This point is perhaps familiar from discussions in other areas of the philosophy of mind.) An acceptable account of the intentionality of emotion, it seems, will have to allow for emotions to sometimes be directed at things other than their causes.

Finally, it is hard to imagine how mere feelings of physiological changes could ever be subject to rational or moral evaluation. On at least a simple feeling theory, like the one that James defends, emotions are apparently no different, from a rational or moral perspective, than feelings of other physiological phenomena—e.g., the feeling of hunger, the feeling of a headache, a sunburn, etc. Without making any reference at all to the conceptual ties that exist between emotions and the sorts of things that typically elicit them, there doesn't seem to be any grounds left upon which James might claim that we ought or ought not have certain emotions in certain circumstances. If fear, for instance, is really *nothing more than* the feeling of such things as an increased heart rate, dryness of the mouth, and trembling limbs, then what makes it the case that I *ought* to be afraid of poisonous snakes and not, say, images of poisonous snakes on the pages of a magazine? How could I ever be morally blameworthy or praiseworthy for feeling (or failing to feel) certain changes to my body? If emotions are nothing more than feelings of physiological changes, I could be no more blameworthy for my failure to feel shame upon humiliating another person than for also failing to feeling hungry, or cold, or exhausted.

So by my count, James' feeling theory can account for only one of the four ordinary features of emotion that I describe above. Emotions are typically accompanied by feelings, and James can explain why this is the case. But he cannot explain the sorts of conceptual ties that exist between the emotions and their elicitors; he cannot adequately account for the intentionality of emotion; and he cannot explain why emotions are sometimes subject to rational and moral evaluation. Perhaps a more sophisticated feeling theory will do the trick. In the next section, I consider just such a theory.

2. Emotions as embodied appraisals

In his book *Gut Reactions* (2004), Jesse Prinz attempts to vindicate James by modifying the basic feeling theory in a way that is supposed to allow it to avoid the problems I just raised for it. So according to Prinz's theory, emotions are still *embodied* in the sense that they are essentially feelings of physiological changes. But, Prinz adds, emotions are also mental representations, and what they represent are "organism-environment relations that bear upon wellbeing" (following Lazarus,⁹ Prinz calls these relations "core relational themes"). When an organism is in danger, it bears a certain sort of relation to its environment; and it is this relation that fear represents. In virtue of their capacity to represent core relational themes like danger, Prinz thinks, emotions count as *appraisals*. So the main innovation of Prinz's embodied appraisal theory is supposed to be that it allows for a sense in which emotions can be *directed at* things *outside* of the body (e.g., danger), while still being identified with *feelings* of changes *in* the body.

Why think that emotions are mental representations? On the account of representation that Prinz prefers, mental representations are mental states that satisfy two

⁹ Lazarus 1991.

conditions: first, they must carry information; and second, it must be possible for the states to be erroneous in some sense.¹⁰ How does a mental state carry information, and what sort of information will it carry? Basically, whenever some thing *x* is reliably caused by some other thing *y*, Prinz thinks, *x* will carry information about *y*. Smoke carries information because it has a reliable cause, and it carries information about *fire* because fire is its reliable cause. If some mental state *M*, then, has a reliable cause *C*, then *M* will carry information about *C*. But this is not yet to say that *M represents C*, for it remains to be determined whether or not *M* satisfies the second condition for mental representation—i.e., the possibility of error. Smoke only carries information about fire, and does not represent fire, because even in cases in which there is smoke but no fire, it does not make sense to say that the smoke is *mistaken*.

What makes it possible for representations to be erroneous, says Prinz, is the fact that they have been “set up” to carry information, i.e., they have the function of carrying the information that they carry. Smoke was not “set up” to carry information about fire; it just does. Our concepts, on the other hand, do seem to have the *function* of carrying certain information.

A dog concept is different [than something like smoke]. It is set up—that is, learned—in order to reliably respond to dogs. A dog concept is a mental state that is reliably caused by dogs *and* was acquired for that purpose. In the simplest cases, this just means that the mental state was initially formed as a result of dog encounters, rather than encounters with something else.¹¹

¹⁰ This is a theory of mental representation that Prinz attributes to Dretske (1981, 1986), but Prinz himself has also defended a version of the theory (2000, 2002).

¹¹ Prinz 2004: 53.

Our dog concept *represents* dogs, then, because the concept was “set up to be set off” by dogs. This is its function. But importantly, our dog concept can *malfunction*, by being elicited by non-dogs (e.g., wolves, foxes, etc.). In such cases, the concept will be applied erroneously. So it is this feature of mental states—i.e., their being set up to be set off by certain things—that allows them to be not just information carriers, but representations.

How is this account of mental representation supposed to provide Prinz with solutions to the problems I raised earlier for James’ feeling theory? According to Prinz, emotions are mental representations because they have been set up (by evolution and learning) to be set off by certain things. Suppose this is the case. Now we must ask, what have they been set up to be set off *by*? That is, what do emotions represent? For an answer, Prinz suggests we reflect upon the conditions that typically elicit emotions: fear typically occurs in circumstances of danger; anger typically occurs in circumstances in which someone has committed an offense; grief typically occurs in circumstances of loss; etc. All of these circumstances involve some kind of organism-environment relation that bears upon wellbeing—i.e., core relational themes. The fact that emotional experiences are patterned in this way, he thinks, is evidence that they have been set up to be set off by—i.e., that they represent—these core relational themes. Furthermore, it is certainly plausible that evolution would be interested in equipping us with mechanisms for detecting such organism-environment relations as *being in danger*. Surely, having a mechanism for detecting dangers is advantageous for the purpose of survival—whereas, for instance, being equipped with a mechanism for detecting things like trembling limbs and dryness of the mouth does not appear to be advantageous in this way. So, Prinz

concludes, “emotions are reliably caused by both bodily changes and core relational themes, but they seem to *have the function of* detecting only the latter.”¹²

Does Prinz’s embodied appraisal theory fare any better than James’ feeling theory when it comes to accounting for the four features of emotion described earlier? I think it does. For starters, if James accounts for the fact that emotions are closely associated with feelings, then so does Prinz. After all, Prinz agrees with James in conceiving of emotions as feelings of physiological changes. So Prinz’s theory can apparently fare no *worse* than James’ theory. Further, Prinz’s account of mental representation allows him to explain the sorts of conceptual ties described above between emotions and their eliciting conditions. On James’ theory, you will recall, it is as if danger has nothing to do with *what fear is*—fear is just the feeling of a particular set of physiological changes, *regardless* of what happens to elicit those changes. But according to Prinz, fear is the feeling of a set of physiological changes *set up to be elicited by dangers*. Let’s allow, then, that Prinz can account both for the fact that emotions are paradigmatically felt, and also for the sorts of patterns that exist between emotions and their eliciting conditions. So Prinz does a better job than James of accounting for the ordinary features of emotion. But I nonetheless think that Prinz’s theory fails adequately to account both for the intentionality of emotion and also for the fact that emotions are sometimes subject to rational and moral evaluation.

Before I explain why I think Prinz does not adequately account for the intentionality of emotion, let me briefly say two more things about the way in which he attempts to do so. First, Prinz thinks that emotions actually have two kinds of objects:

¹² Prinz 2004: 67; italics mine.

A *formal object* is the property in virtue of which an event elicits an emotion, and a *particular object* is the event itself. The death of a child can be a particular object of one's sadness, but it causes sadness in virtue of being a loss. Being a loss is the formal object of sadness. Emotions represent their formal objects, not their particular objects.¹³

If emotions only ever represent their formal objects, though, how do particular objects get represented? According to Prinz, an emotion's particular object is represented by some non-emotional mental state that, in some sense, accompanies the emotion.

If I am sad about the death of a child, I have one mental representation [e.g., a belief] that corresponds to the child's death and another, my sadness, that corresponds to there having been a loss. Together, we can think of these as constituting a complex representation that means the child's death has been a loss to me. We might think of the compound as meaning something like: a child has died, *and* what a loss!¹⁴

And second, like James, Prinz apparently believes that the relationship between an emotion and its object—whether its formal or particular object—is a *causal* relationship. “The conditions (real or imagined) that elicit an emotion,” he writes, “can be referred to as its objects.”¹⁵ He is even more explicit a few lines later: “Saying that my sadness is *about* the death does not mean that my sadness represents the death; rather it means that the death is what *caused* me to become sad.”¹⁶

As we saw earlier, though, *however* it is that emotions take their objects, it is apparently not by being *caused* by them. In the earlier example, my anger was directed at a slandering co-worker who may or may not even exist. Interestingly, Prinz himself appears to countenance cases of non-existent and imaginary objects of emotion—“The

¹³ Prinz 2004: 62. Prinz borrows this distinction from Kenny 1963.

¹⁴ Prinz 2004: 62-63.

¹⁵ Prinz 2004: 62.

¹⁶ Prinz 2004: 62, second italics mine.

conditions (*real or imagined*) that elicit an emotion can be referred to as its objects.” But it is unclear how an imaginary object—i.e., the *object*, not the *imagining* of it—can cause a person to experience an emotion. Again, in the example, I am angry *at John*. *He* is the object of my anger. Surely, though, John could not have caused me to become angry.

Furthermore, even if Prinz’s analysis seems correct in some cases, he owes us more by way of an explanation of why only *one* cause of an emotion, and not *others*, turns out to be the emotion’s object. It is not enough to say that the death caused me to be sad, for presumably, my sadness could have been caused by more than *just* the death. Maybe this particular death only causes me to become sad because of how close I’ve recently grown to a relative of the deceased. Had I not recently grown so close to this person, I would not be at all saddened by the death. So what is the cause of the sadness? The death? Or my close relationship to the relative? Both, it seems. But of course, my sadness is only about one of these things. If emotions take their objects in virtue of being caused by them, however, one wonders why we do not also say that my sadness is about the close relationship. Maybe I had too much coffee to drink on the day that Susan reports my non-existent colleague’s slanderous behavior, and the effect of the coffee also has a role to play in eliciting my anger. So why is my anger only *at John*, and not also at the coffee (or its effects), or Susan’s reporting, or the elaborate hoax to get me to believe that John exists in the first place? If not for any of these other things, I wouldn’t be angry. But none is the object of my anger. Prinz offers us no way to explain why, on his causal account of the way in which emotions take their objects, emotions only take some

causes as their objects and not others.¹⁷ All the more reason, I submit, for thinking that the *object-of* and *cause-of* relations operate fundamentally differently when it comes to emotion.

Can Prinz account for the fact that emotions are sometimes subject to evaluation? In one sense, yes; but in other important senses, no. According to Prinz, emotions can be “misapplied,” just as, for instance, concepts can be misapplied. I might see a wolf and think to myself, “That *dog* is huge.” Here a particular mental representation—the dog concept—is being applied erroneously. Likewise, if fear has been “set up to be set off by” dangers, then whenever I am afraid of a non-danger, my fear is similarly erroneous. This is the sense in which Prinz can account for emotions being subject to evaluation: for any emotion, it *ought* always and only to be caused by whatever it has been set up to be set off by; and whenever it is not, the emotion is in this sense mistaken. This, indeed, is an improvement upon James’ theory, which apparently could not account for *any* sense in which emotions are subject to evaluation. But there are still cases in which our emotions are subject to types of evaluation for which Prinz apparently cannot account. First of all, he cannot account for any sense in which emotions are subject to moral evaluation. When I say of someone else, “You ought to be ashamed of yourself,” the ‘ought’ here is clearly a *moral* ought. What I mean to be saying is that, in this circumstance, a lack of shame constitutes a moral failing. But nothing about Prinz’s account of emotion explains why this should be the case.

¹⁷ Of course, he could say that the answer lies in whatever non-emotional state has come to be “attached” to the emotion—e.g., in the case of the death, it is a thought *about the death*, and not a thought about my close relationship with the relative. But this would only be to push the problem back a step: Why is it that *that* thought, and not another, comes to be attached in this way to the emotion?

There also appear to be cases of emotions that are “erroneous” in the sense that Prinz describes, but are also evaluable along a further dimension for which he cannot account. Consider two people who are deathly afraid of flying. One, my mother-in-law Martha, knows that flying is a relatively very safe way to travel. She was a flight attendant for many years, and her husband was a pilot for his entire professional life. She has memories of trying to talk terrified passengers out of their own fear, and to this day, could give lectures on the safety of flying relative to other forms of travel. But for all that, she is now too scared to step foot on a plane. The other person, Stewart, shares Martha’s fear of flying, but also believes (falsely) that flying is one of the most dangerous ways to travel. Notice, both fears are erroneous in the sense that Prinz describes—i.e., they’re both apparently being “set off” by non-dangers. But there is something else that is going wrong, something else that is out of order, in the case of Martha’s fear, but not in the case of Stewart’s. There is a sense in which Martha’s fear *shouldn’t be there*, but Stewart’s *should*, even though they’re both erroneous in Prinz’s sense. And it is difficult to see how his view could account for this further kind of error. Irrational fear is not simply fear of a non-danger; rather, for me to have an irrational fear is for me to be afraid of something *I do not see as a danger*. Cases like my mother-in-law’s fear of flying suggest that our emotions owe something not just to *the way things are*—as Prinz explains—but also to *the way we see things*.

The embodied appraisal theory of emotion is no doubt an improvement upon James’ feeling theory. Prinz can still account for the fact that emotions are closely associated with feelings, and it also appears that he can accommodate the clear conceptual ties between emotions and their eliciting conditions. But I do not think that he

accounts for the intentionality of emotion in the right way. And I think we should prefer a theory that better accommodates the various dimensions along which our emotions seem subject to rational and moral evaluation.

3. Emotions as judgments¹⁸

Switching gears a bit, I'll now briefly consider a theory according to which emotions are not anything physiological. Rather, on this theory, emotions are judgments of a certain sort. Here is Robert Solomon stating the theory:

“I am angry at John for taking (“stealing” begs the question) my car” entails that I believe that John has somehow wronged me. (This must be true even if, *all things considered*, I also believe that John was justified in taking my car.) The (moral) judgment entailed by my anger is not a judgment *about* my anger [...] My anger *is* that judgment. If I do not believe that I have somehow been wronged, I cannot be angry (though I might be upset, or sad). Similarly, if I cannot praise my lover, I cannot be in love (though I might want her or need her, which, traditional wisdom aside, is entirely different). If I do not find my situation awkward, I cannot be ashamed or embarrassed. If I do not judge that I have suffered a loss, I cannot be sad or jealous. I am not sure whether all emotions entail judgments [...] But emotions in general do appear to require this feature: to have an emotion is to hold a normative judgment about one's situation.¹⁹

Right off the bat, we can make a few observations about Solomon's theory. First of all, it probably ought to be tweaked a bit in order to allow for cases of genuine anger even when the subject does not judge that *he* is the one who has been wronged. My anger at some child abuser I hear about on the news surely does not entail a judgment to the effect that the child abuser has wronged *me*, but rather only that he has wronged *someone*—or

¹⁸ For the record, in order to discuss judgment theories of emotion, I'll have to temporarily suspend my insistence upon only using 'judgment' to refer to a certain sort of psychological process (see §3 of chapter I). Emotion theorists like Solomon clearly use 'judgment' in the belief-sense, so that is the sense in which I will use the term in this section.

¹⁹ Solomon 1980: 257-258.

perhaps even more broadly, that he has committed some kind of offense. And of course, something similar will likely have to be said about many other emotions—I can be afraid even if *I* am not the one in danger; I can grieve even if *I* am not the one who as suffered a loss; etc.

Second, Solomon’s theory will apparently have no trouble accounting for the sorts of conceptual ties that exist between emotions and their eliciting conditions, since the eliciting conditions typical of a particular emotion will presumably be referenced in the very judgment to which the emotion is supposed to be identical. Fear is related to danger, Solomon will say, because to fear *x* is just to judge that *x* is a danger. Anger is related to offense, because to be angry at *x* is just to judge that *x* has committed an offense. And so on and so forth.

And third, this sort of theory is clearly better suited to account for the intentionality of emotion than the other two. According to Solomon, emotions are just judgments, and judgments are obviously among the sorts of things that can be intentional. This theory can explain not only why it is, for instance, that sadness or grief is about loss-in-general, but also how it is that particular emotions have the particular objects that they do. When I grieve my grandfather’s death, my grief is not just about loss-in-general, but about *that* loss in particular. For my grief in this case is my judgment that *my grandfather’s death is a loss*. In fact, the apparent ease with which a theory like Solomon’s can account for the intentionality of emotion has led many to believe that the best theory of emotion will have to be one that brings some kind of “cognitive” element—at least, something more cognitive than a mere feeling, like a belief, judgment,

etc.—into the analysis of what an emotion is.²⁰ Obviously, Solomon does this by saying that emotions are *constituted by*, if they are not *identical to*, judgments of a certain kind; but as we'll see later, this is certainly not our only option.

How about the other features of emotion? Can Solomon account for various ways in which our emotions are sometimes subject to evaluation? He thinks he can. If emotions are just a species of judgment, then they are subject to all the same criticisms to which judgments are typically subject. This includes criticisms made on the basis of inaccuracy: to be afraid of a non-danger is to make the same sort of mistake one makes in judging that a non-dog is a dog. So Solomon can apparently account for the same sense in which emotions are subject to evaluation for which Prinz could account. But Solomon thinks he can also account for the moral responsibility that we sometimes attach to people for their emotions.

[I]t always makes sense, at least (as it does not, for example, for headaches, heart attacks, and hormones) to praise or blame a person, not just for contributing to the situation that caused the emotion but, [...] for having the emotion itself, as one blames a person for bigotry, for example, or praises them for courage.²¹

This is because, as Solomon argues, rather than *happening to us*, emotions are things that we *do*. Emotions are judgments, and judgments are things we *choose to make*. They are actions. And so we are every bit as much to blame for feeling contempt or disgust toward someone of a different race as we are for using a racial slur.

²⁰ It is worth noting here that the term 'cognitive' is notoriously difficult to define, perhaps especially in discussions of the nature of emotion. This issue will reappear later in the dissertation. For more discussion, see Prinz 2004: 41ff and Moors 2010: 1-2.

²¹ Solomon 1980: 276.

On the one hand, Solomon certainly deserves credit for establishing some sense in which people can be morally responsible for their emotions. Prinz, apparently, could not do this.²² But on the other hand, it is surely false that emotions are always *chosen* by us in the sense that Solomon thinks they are. Suppose Tom was raised in a racist household, conditioned over years to feel contempt and disgust toward black people. As an adult, however, Tom has since disavowed the racism of his upbringing. He has friends and co-workers who are black, and he now sincerely judges that black people are his equals in every relevant sense. But for all that, there are times at which he cannot help himself. When his daughter begins dating a young black man, for instance, he finds himself overcome by the sorts of feelings he once had toward black people. He is ashamed by this, but finds that there is little he can do to change his feelings toward this young man.

It is no doubt regrettable that Tom feels this way toward his daughter's boyfriend, but given all that we know about him, it cannot be right to say that he is *choosing* to feel contempt and disgust toward the young man. This is the difference between Tom and his relatives (who, presumably, would feel the same toward the young man), and it is a difference that is surely relevant to our evaluation of these people and their emotions. Given the circumstances, I'm certainly inclined to blame the relatives for their emotions (and maybe also for Tom's); but I'm more inclined to pity Tom than to blame him. At the very least, he is surely *less* blameworthy than his relatives, despite having the same emotions. If, however, as Solomon argues, "emotions are rational and purposive rather than irrational and disruptive," and "we choose an emotion much as we choose a course of action," it is difficult to see how we could really draw this sort of distinction between

²² At least, not without an account of moral responsibility according to which one can be morally blameworthy for failing to be caused to have certain feelings.

Tom and his racist relatives.²³ So while I think Solomon has gone some way toward adequately accounting for the various ways in which we hold people responsible for their emotions—certainly more so than Prinz—we still might hope for an account that allows us to distinguish between emotions that are chosen, or under our control, and emotions that are not.

And finally, Solomon's theory suffers for its inability to account for the fact that it is characteristic of emotions that they be accompanied by feelings of various sorts. He presents the theory as a kind of reaction to the prevailing Jamesian view that emotions are "physiological disturbances." Against such a view, he argues, "One can be angry without feeling angry: one can be angry for three days or five years and not feel anything identifiable as a feeling of anger continuously through the prolonged period."²⁴ On the basis of cases like this, Solomon concludes, "[A]n emotion is never simply a feeling, not even a feeling plus anything."²⁵ And as we've seen, he ends up with an analysis of emotion on which feelings are nowhere relevant. As far as Solomon's theory is concerned, someone completely incapable of experiencing any of the feelings we ordinarily associate with emotions—think of someone like Data from *Star Trek*—could nonetheless have every bit as much of the emotional life as anyone else. That is, as long

²³ There is a deeper problem that Solomon faces here, one that may actually entail the falsity of his theory. If emotions *just are* judgments, then it shouldn't be possible for emotions and their corresponding judgments to part ways, as it were. But in some cases, this seems to be exactly what is going on. In the above example, Tom is disgusted by someone he judges to be his equal. In the earlier example, my mother-in-law is afraid of something she judges to be relatively safe. In response to such cases, it seems Solomon will have to insist either that these people do not really make the judgments they appear to make (e.g., that my mother-in-law does not *really* judge flying to be safe), or else that they are not really experiencing the emotions they appear to be experiencing (e.g., that my mother-in-law is not *really* afraid of flying). I think neither of these responses is plausible—and so, I think he is wrong to *identify* emotions with judgments—but I'll not pursue the matter any further here.

²⁴ Solomon 1980: 254.

²⁵ Solomon 1980: 254.

as this unfeeling person has whatever is required for judging things to be dangers, offenses, losses, etc.

This looks to me like an *over*-reaction to the Jamesian view. James may go too far in identifying emotions with feelings, but Solomon goes similarly too far in severing feelings from emotion altogether. Even if he is right that emotions can sometimes be unfelt,²⁶ such cases are atypical, and nothing about Solomon's view explains why this is the case. In fact, feelings are apparently so integral to our emotional lives that we have developed a habit of talking about emotions as if they were nothing more than feelings. When I want to know how some event affected my friend emotionally, I ask him, "How did that make you *feel*?" We say things like, "That makes me feel angry," or, "I felt so afraid." If emotions are just judgments, though, it is hard to imagine what an expression like "feeling afraid" could even mean. Feeling as if (I've judged that) something is a danger? What does that mean? The fact of the matter is that there are feelings characteristic of the emotions, and we should prefer a theory of emotion that accommodates this fact.

I think Solomon's theory is a step in the right direction from Prinz's. Solomon can account both for the conceptual ties between emotions and their elicitors, and also for the intentionality of emotion (at both the general and particular levels). He can also account for a sense in which emotions are subject to both rational and moral evaluation, though I think there are problems with the way he does this. So perhaps we should grant that he's captured three of the four ordinary features of emotion described in my introduction. Each section has gotten us closer and closer to a satisfactory account of

²⁶ I'll argue below that emotions actually *cannot* be (completely) unfelt.

emotion, and in the next section, I'll attempt to bring the discussion home with an account that captures all of the ordinary features of emotion that I have been discussing.

4. My account of emotion

In this section, I develop an account of emotion according to which emotions are complex states, consisting of three things: construals, concerns, and feelings. My account of emotion owes much to the account defended by Robert Roberts (1988, 2003), according to which emotions are “concern-based construals.” In fact, construals are central to my own account as well, so I'll happily draw upon relevant parts of his work in order to expound my view. In the end, it may turn out that the differences between our two accounts are negligible at best, but I'll nonetheless argue that my own account does a better job than Roberts' account of explaining certain cases.

Following Roberts, I understand a *construal* as “a mental event or state in which one thing is grasped in terms of something else.”²⁷ This definition is intentionally imprecise, as the set of events or states that can count as construals is actually quite varied. For instance, the sort of “grasping” that I have in mind can take many forms: perceiving, thinking of, imagining, judging, etc. And the set of things in terms of which other things can be grasped is similarly diverse, including perceptions, images, thoughts, concepts, etc. In *Philosophical Investigations* (1953), Wittgenstein considers a case in which “I contemplate a face, and then suddenly notice its likeness to another. I see that it has not changed; and yet I see it differently.”²⁸ Here, for instance, we might say that Wittgenstein is grasping the first face in terms of the second—i.e., he has gone from

²⁷ Roberts 1988: 190.

²⁸ Wittgenstein 1953: 193.

simply *seeing* a face, to seeing the same face *as* (similar to) another. We can also see faces in terms of concepts, e.g., seeing the face *as rugged*, or *as kindly*, etc. Construals, then, are always cases of *seeing-as*, but the “seeing” is not always straightforwardly perceptual.

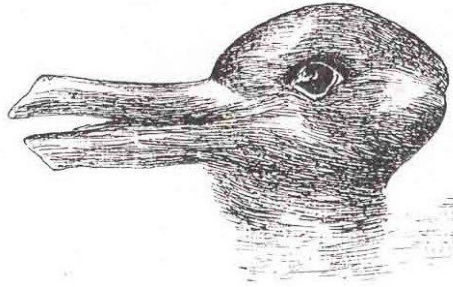
I can imagine my living room in terms of furniture in the store, which I am presently perceiving; or in terms of either the image or the thought of my parents’ living room, or in terms of the concept *grandiose* or *well-coordinated*; or I can think of myself (which is quite different from imagining myself or perceiving myself) in respect of likeness to my father, whom I am presently perceiving; or in terms of an image, such as coming sweating and triumphant across the finish line; or in terms of a concept, like *intelligent* or *moody*.²⁹

So construals are not always perceptual, but as Roberts explains, they very often “have an immediacy reminiscent of sense perception. They are impressions, ways things appear to the subject.”³⁰ Just as soon as I see a man raise his auction paddle, I see him *as bidding*; just as soon as I see a gray cloud, I see it *as a rain cloud*; and so on and so forth.

Construals are sometimes under our control, and sometimes not. The coffee cup before me appears white, and barring some sort of intervention (putting on colored lenses, etc.), there’s little that I can do about that. However, I might construe the cup in any number of ways: as a gift from my wife, as in need of cleaning, as half-empty, as half-full, etc. And in many cases, I will be able to switch back and forth between these construals voluntarily. Consider the famous duck-rabbit figure:

²⁹ Roberts 1988: 190.

³⁰ Roberts 2003: 75.



Initially, I might only be able to see this as a duck. But once I get the hang of seeing the rabbit as well, I may soon be able to switch back and forth at my leisure. Of course, I am no more in control of the visual input in this case than I was in the case of my white coffee cup, i.e., I am not in control of *what I see*. But I am—or can be—in control of *the way I see it*. Sometimes, however, a particular way of seeing something can have a kind of grip on you, so that you cannot help but see it in that way. And this need not be because you haven't yet learned how to see it differently, as in the case of a visual illusion like the duck-rabbit. For example, it can sometimes be incredibly difficult to see a person as anything *but* a political opponent, an unfaithful lover, or a character that he or she once played on television.

Some construals are unconscious. If you've ever found yourself consistently referring to a person by someone else's name, for instance, it may be that you are unconsciously construing the one in terms of the other. Here is another example:

I am reading a journal article in which a quotation appears at the end of one page and extends by three words onto the next. I turn the page and can't find the end of the quotation. I turn back again to see whether the quotation has in fact ended on the preceding page. No. On further examination of the succeeding page I see the three words hanging there isolated at the top, and realize that I have been construing them as a header, with the result that I did not consciously see them at all.³¹

³¹ Roberts 2003: 72.

Other examples are more serious. As in the case of Tom above, if you were raised in a racist household, for instance, you may find that you respond negatively to people of a different race even if you have since disavowed the racism around which you were brought up. And this is because, unconsciously, you still construe people of a different race in the terms in which they were once characterized by those around you. Interestingly, unconscious construals may also lie behind a phenomenon highlighted by Prinz in his critique of cognitive theories of emotion like Solomon's. Prinz cites studies that suggest that we can alter our emotions by simply making the facial expression characteristic of a particular emotion (e.g., we can make ourselves happier by smiling, angrier by scowling, etc.).³² According to Prinz, this is evidence that emotions can be activated without the mediation of an intentional attitude like judgment. However, if construals can be unconscious, there is another interpretation available. Here is Roberts:

When facing that fearsome interviewer, try sitting up straight, leaning forward in a slightly assertive bodily attitude, looking her in the eye, and talking in an even voice. Doing so will make you appear (to the interviewer, but more importantly, to yourself) to be in control of the situation. You help out the prescribed construal by making the situation more plausibly construable in its terms.³³

Perhaps something similar is taking place in the case of facial expressions: In making a certain facial expression, we help ourselves to (unconsciously) see a particular circumstance as one that warrants whatever emotion is relevant to the expression. As far as I can tell, Prinz does not seriously consider this possibility.

How is all of this talk about construals relevant to emotion? According to my preferred account of emotion, emotions essentially involve construals, so one cannot have

³² Strack, Martin, and Stepper 1988; Zajonc, Murphy, and Inglehart 1989.

³³ Roberts 1988: 194.

an emotion without construing something in some way. But not every construal is involved in an emotion. The difference between a construal that *is* involved in an emotion and one that *is not* has to do with the ways in which the two construals relate to our *concerns*. According to my account, if Smith and Jones both construe themselves as having endured a loss of some sort (e.g., the death of a relative), but only Smith had any concern for the thing lost (e.g., Smith, but not Jones, loved the deceased), then only Smith will be sad. Likewise, if Smith and Jones both construe themselves as having been mocked by Davis in front of their colleagues, but only Jones is concerned about his professional reputation, then only Jones will feel embarrassed, or angry with Davis. Suppose I'm given a giant, gaudy lawn ornament as a gift, but for whatever reason, can neither refuse it nor keep it hidden. I spend my days and nights hoping and praying that it will be taken from me. When it eventually is stolen, I may construe the thief as having stolen something from me; but even if I do, I may not be angry with him. On the other hand, if he had stolen my car, something for which I have a great deal of concern, I would be furious. So emotions involve construals, but not just any construals—rather, emotions involve only those construals that either clash or accord with our concerns.

Before moving on, let me make two quick notes about the concerns to which construals are supposed to relate in these ways. First, I mean for the term 'concerns' to be understood very broadly. Included will be everything from aesthetic preferences, to loves (of perhaps all kinds), to moral, political, and religious commitments, to desires, goals, etc. And second, the concern in question will have to be of a sufficient magnitude (from the subject's perspective) in order for a construal to give rise to an emotion. For instance, I can construe my coffee cup as empty, and that construal might clash with my

mild desire for more coffee, but it is unlikely that this will result in anything I'd consider to be an *emotional* experience. On the other hand, if I have a deep love for coffee and know that this cup will be my last ever, then my construal of the cup as empty might lead to great sorrow on my part. (I'll have more to say about construals in chapter IV.)

Here is where my account of emotion differs from Roberts' *concern-based construal* account. Roberts thinks that concerns must be *a part of* the construals themselves, i.e., a part of the way we see things. He writes, "[T]he concern enters into the construal so as to characterize the appearance of the object."³⁴ This implies that two people with a different set of concerns with respect to *x* cannot construe *x* in the same way, i.e., they must see *x* differently. But that doesn't seem right. In the above examples, it is not that Smith and Jones see themselves differently—each sees himself as having lost a relative, or as having been mocked by Davis. If Roberts is right, though, it would seem that Smith just *cannot* construe Davis as mocking him, or as threatening his professional reputation, since he lacks the relevant concern. But why think this? Consider another case. There is an activity known as "BASE jumping" that involves jumping off of some fixed object (e.g., buildings, bridges, cliffs, etc.) and quickly deploying a parachute, the canopy of which the jumper holds in his or her hand. It is widely considered to be one of the world's most dangerous recreational activities. Suppose that a BASE jumper and I (not a thrill-seeker by any stretch of the imagination) stand on the edge of Angel Falls in Venezuela, the world's highest waterfall. The thought of jumping off of the 3,212-foot-high waterfall terrifies me. But the BASE jumper is positively giddy with excitement. Roberts' view implies that this BASE

³⁴ Roberts 1988: 192.

jumper and I must be seeing things differently, and to be sure, it is not hard to understand why one would be inclined to think so. After all, we *respond* to the situation in very different ways. But if this thrill-seeker did not see the very same danger that *I* see when construing the situation, presumably, she would not be as excited as she is. After all, if thrill-seekers themselves are to be believed, the sense that one's safety is seriously threatened is precisely what makes it thrilling. According to Roberts' view, however, it seems we'd have to say that she cannot see jumping off the waterfall as a real threat, or else she'd be afraid.³⁵ Perhaps there is a way for Roberts to account for cases like this. For now, though, I maintain that this one difference between our accounts—i.e., that he builds concerns *into* construals, and I do not—represents an advantage of my view: I am able to accommodate cases in which two people see things similarly but respond emotionally in different ways, whereas Roberts must insist that this is impossible.

On the account of emotion that I have been developing, emotions involve (essentially) construals and their relations to our concerns. But there is still one more piece to put in place: *feelings*. While I think James and Prinz are mistaken to *identify* emotions with feelings, I do agree that emotions always *involve* feelings. I do not, however, conceive of the relevant feelings as only perceptions of physiological changes. Unfortunately, surely due in no small part to the influence of James' work, emotion theorists often speak of feelings in this context as if they are essentially bodily in nature. But this does little justice, I think, to the ways in which we ordinarily talk about

³⁵ Indeed, in conversation, Roberts has explained, "She [the thrill-seeker] doesn't see it as a threat, although she is no doubt aware that she is doing something a bit dangerous, even for someone who has her equipment and skill. But it doesn't appear anywhere near as dangerous to her as it does to you." I think this is precisely what we will have to say if we think that concerns are a part of the construals themselves, but I maintain that it does not really do justice to the psychology of the thrill-seeker.

emotional feelings. Of course, many of the feelings involved in an emotion *are* somatic, but what about things like the feeling of down-ness associated with emotions like grief, or the feeling of triumph one experiences in victory, or the awkwardness one feels when traveling abroad and trying to speak a language in which one is far from fluent? These feelings seem to be more psychological than physiological. (Oakley calls them “*psychic feelings*,” and contrasts them with bodily feelings.³⁶) One of the ways in which shame might be felt, for instance, is in an inability to look people in the face (especially those with knowledge of the situation).

What about cases of so-called “unfelt emotions”? Consider a case like the one mentioned by Solomon in which someone is supposed to have been angry at *x* for years, and yet has not felt anything identifiable as the feeling of anger continuously through this period. The assumption is supposed to be that, even at those times at which the subject feels nothing of the anger, he is still angry. In such a case, however, I want to say that the person has not actually *been angry at x for years*, but rather has *been disposed* for years to be angry at *x* because, apparently, his anger at *x* has gone unresolved for years. This is just what we mean when we say things like, “He’s been angry for years.”³⁷ Consider

³⁶ Oakley 1992: 7ff.

³⁷ Here is a related example: Someone who says, “I don’t get angry; I get even.” One might think it plausible that, while such a person may not *feel* angry, she nonetheless *is* angry; and in getting even, she is acting *out of* anger. I have a few things to say about this case. First of all, there are real-life cases of people who perform vengeful actions despite an apparent deficit in relevant emotions—psychopaths, for instance. These people presumably do not experience the sorts of feelings that we associate with certain emotions, but this is because they’re not experiencing the emotions at all—they don’t *feel* angry because they are *not* angry. So the fact that one acts so as to “get even” is no guarantee that one is acting *out of* anger, or any other emotion. Second, in normal cases, when people say things like this, I don’t think they mean to be giving an honest report of their psychological state. Think, for instance, of the saying, “I don’t *think* it; I *know* it.” This is pretty clearly not meant as a *report* of one’s epistemic position with respect to some proposition or state of affairs; rather, it is meant as *an expression of confidence*. Likewise, the saying, “I don’t get angry; I get even” is meant not as a report of one’s psychological state, but rather as an expression of one’s intention to settle a score. At least, this is how I’ve always interpreted someone who

another example: Rachel says that, throughout her entire engagement period, she was anxious about getting married. But of course, on any one of the days during this period, we might have found her, say, typing away at her dissertation, engrossed in research, her upcoming nuptials the furthest thing from her mind. In what sense is Rachel, in that moment, anxious about getting married? She is anxious in the sense of being disposed to anxiety: if I interrupt her work and begin asking about her wedding plans, she'll be anxious, and her anxiety will be accompanied by feelings of various sorts. But at the moment, she is not actually anxious—just disposed to it. Cases of so-called unfelt emotions are always cases of dispositions to feel emotions. Emotions, I maintain, are always felt in *some* way.

So, in my view, emotions are felt harmonies or disharmonies between the ways in which we see things (construals) and our concerns, and the sort of “feeling” I have in mind can involve any combination of physiological and psychological feelings. To be clear, on this view, emotions are *neither* construals, *nor* concerns, *nor* feelings; rather, emotions are complexes consisting of all three, and the relations they bear to each other.

How does this view accommodate the four ordinary features of emotion that I've been discussing throughout the chapter? First, obviously, the view can accommodate the fact that emotions typically involve feelings of various sorts—on this view, as with the views of James and Prinz, those feelings are a part of the emotion itself. Second, it nicely captures the connections between things like fear and danger, anger and offense, grief and loss, etc., because the latter items in these pairings will be among the terms in which

says something like this. And finally, if this person really *is* angry, then I'd simply insist that she *is* experiencing some sort of emotional feelings, even if those feelings are currently being swamped, phenomenologically, by her desire to get even.

the subject of the emotion construes things. For instance, fear will typically involve a construal of something as a danger. And third, it accounts for the intentionality of emotion in virtue of the fact that construals are themselves intentional states. The object of my emotion will always be whatever is the object of the relevant construal. The snake is the object of my fear because it is the snake that I am seeing as dangerous. Davis is the object of my anger because it is Davis that I construe as having committed some offense against me. I am the object of my shame because I see myself as having done something terrible.

In these ways, construals operate on my account a bit like judgments operate on Solomon's. But construals are importantly different from judgments, in that *construing* x as F does not entail that the construing subject actually *believes* that x is F —whereas one cannot *judge* that x is F without also believing that x is F . This is important because it allows me to avoid a problem raised above for Solomon's theory (see n. 23): the problem of accounting for the occasional dissociation between our emotions, on the one hand, and our beliefs, judgments, and the like, on the other. If fear of x is a judgment that x is a danger, then one can apparently not fear something that one judges to be safe. But this is precisely what's happening in the case of my mother-in-law's fear of flying. On my view, cases of dissociation are easily accounted for, since I can judge that x is F and still *see* x as *not- F* . My mother-in-law may sincerely judge that flying is among the safest ways to travel, and yet still *construe* flying (perhaps unconsciously) as a danger. Likewise, Tom may now judge that black people are his equals and still see his daughter's boyfriend as having one or another of the qualities he was raised to associate with black people.

This brings me to my explanation of the ways in which emotions are sometimes subject to rational and moral evaluation. To begin with, other things being equal, we ought to see things as they are. Now, in some cases, there will be more than one way to do this. For instance, in the case of the duck-rabbit, it is no more (or less) appropriate to see it as a duck than it is to see it as a rabbit. To see it as a copy of Herman Melville's *Moby Dick*, however, is to see it as something that it isn't—and so, to construe it inaccurately. Construals, I submit, are subject to norms of accuracy. When I construe the coffee cup as a gift from someone *other* than my wife; whenever someone construes a wolf as a huge *dog*; and when my mother-in-law construes flying as a serious *danger*—we are all violating a general norm that we ought to see things as they are.³⁸ This is just one sense in which emotions, on my view, can be subject to evaluation, and it is the sense for which *both* Prinz and Solomon could account. (To return to an earlier example, on my account of emotion, this is why it is right to fear a ferocious bear and wrong to fear Yogi bear—because only the former, and not the latter, is accurately construed as a danger.)

Now recall the earlier contrast between my mother-in-law Martha and Stewart, a person who shares her fear of flying, but also sincerely believes that flying is a real danger. In response to this case, I suggested that there are actually two kinds of mistakes being made here, one that both Martha *and* Stewart are making, and another that only

³⁸ Of course, there are contexts in which these norms do not apply, i.e., contexts in which violating these norms does not count as any kind of mistake or wrongdoing on our part. In order to better understand someone else's perspective, for instance, I might try "seeing things through her eyes," i.e., construing things as (I think) she construes them. And this might involve construing a thing as something that it isn't—e.g., construing flying as a relatively dangerous way to travel. In such a case, it surely wouldn't be right to say that I'm guilty of some sort of error.

Martha is making. (And Prinz, I argued, cannot account for the second of these mistakes.) The mistake that both are making is the one described in the last paragraph: each fails to see flying *as it is*, namely, a safe way to travel. The mistake that only Martha makes, however, is different. Construals owe something not only to the way things are, but also to our beliefs or judgments about the way things are. If I believe that x is F , then I have a defeasible reason for *seeing x as F* (or at least, a defeasible reason for *not seeing x as not- F*). Martha's belief that flying is relatively safe gives her a reason for seeing it that way, whether it is in fact safe or not. If she construes flying as dangerous, then, she either does so for some other reason, or else she does so *irrationally*. Likewise, Stewart's belief that flying is dangerous gives him a defeasible reason for seeing flying as dangerous, but it will presumably be a pretty weak reason, given that both construals and beliefs are subject to norms of accuracy.

Finally, my account of emotion nicely explains why it is that our emotions are sometimes subject to *moral* evaluation as well. And it is not because emotions are always choices, or actions, as Solomon controversially argued. Rather, it is because there are some concerns that we ought or ought not have. If you have a desire to see other people humiliated, then you will likely feel delighted or amused whenever you construe a person as such. But if you have this desire, there is something wrong with your character, a blemish brought to light by your amusement. And it is for this reason that, upon noticing that you are amused by another's humiliation, I might say, "That's awful of you. You shouldn't feel that way." Likewise, a failure to feel indignation when another treats you unfairly may be rooted in an unfortunate—and, in some sense, blameworthy—

lack of self-respect. And so, when I notice your lack of indignation, I might say, “What’s wrong with you? You should be indignant!”

Solomon is mistaken to think that emotions are always choices, or actions, but it cannot be denied that there are similarities between emotions and actions. And just as we often coordinate the degree to which we hold people morally responsible for their actions with the degree to which we think they are *in control of* those actions, the same can be said of the emotions. Perhaps in most cases, people are sufficiently in control of their emotions—they’ve chosen to construe things in a certain way, and the relevant concerns are ones with which they fully identify, etc.—and so, they are rightly held fully responsible for them. But as we’ve seen, this is not always the case. Maybe, after saying, “What’s wrong with you? You should be indignant!” I discover that the person’s lack of self-respect is rooted in years of psychological manipulation and abuse. In such a case, it may still be the case that she ought to feel indignation, and I may still wish that she did, but I can hardly blame her for not feeling indignant. In the case of Tom from earlier, it is certainly the case that he ought not feel the way he does toward his daughter’s boyfriend, and we will certainly wish that he didn’t. But we are at least less inclined to blame him for his emotions than we are to blame his racist relatives for theirs. He is trying to see people differently; he is trying to have the right sorts of concerns. His relatives are not.

5. Why this has been important

Judging by the size and scope of historical and contemporary treatises on emotion, I think it is safe to assume that there is quite a bit more that can be said about the nature of

emotion. In this chapter, I have only attempted to develop a rough sketch of what an emotion is, one that I hope to have shown does a better job of accounting for ordinary features of emotion than a few of the more prominent theories on offer. I think an account of the sort that I defend has a number of other explanatory fruits as well. For instance, conceiving of emotions in the way that I do allows us to explain the fact that our emotions sometimes *reveal our concerns*, both to others and even to ourselves. Sometimes, we do not realize how much a thing matters to us until we respond to it emotionally—e.g., “I didn’t realize how much I loved her until she left me and I grieved for months.” This is easy to explain if we allow that emotions always involve concerns, even if they are sometimes concerns of which we are not aware. Further, the relations that may exist between construals and concerns might help to explain the *positive* and *negative* valences that emotions typically have. Emotions like anger, fear, disgust, and grief are negative because they all involve construals that clash with our concerns, whereas emotions like pride, joy, and gratitude are positive because they all involve construals that accord with our concerns.³⁹

So I will now proceed on the assumption that emotions are roughly as I have characterized them in this chapter. Why has this been important? As I mentioned in chapter 1, the purpose of this dissertation is to give an account of the role that emotion

³⁹ There are more explanatory fruits that I think my account of emotion can offer, but due to limitations of time and space, I just cannot pursue them at any length here. For instance, conceiving of emotions in the way that I do may help explain phenomena associated with *emotional maturity*. What is it that distinguishes the emotionally mature from the emotionally immature? The emotionally mature person is not only more accurate in the ways in which she construes things (something that likely comes with a sufficiently wide range of experiences, etc.), but also has more of the concerns that a person ought to have. My account may also explain why the emotional repertoire of adult humans is typically more complex and sophisticated than that of children and non-human animals: because (a) the greater cognitive sophistication of human adults likely allows for a wider repertoire of possible construals, and (b) over time, adult humans are likely to develop a much larger set of concerns.

plays in moral judgment. So now that we have some idea of what emotion is, we can ask: What sort of role does *that* play in moral judgment? This question, though, actually has two meanings, and the picture that I've painted in this chapter of the nature of emotion will be relevant to both.

Suppose a friend tells you that she landed a role in a play, and you ask, "What sort of role?" If she responds by saying, "The lead," then she may or may not have answered your question, depending upon how the question was meant. If by "What sort of role?" you meant to be asking a question about the *significance* of the role—i.e., about how major or minor the role is, compared to others—then she will certainly have answered your question. But if you meant to be asking about the *nature* of the role—i.e., about the character she'll be playing, its background, relationships, and so forth—then her response will not have answered your question *at all* (that is, assuming you have no prior knowledge of what play it is). So when we ask the "What sort of role?" question, we might be asking a question about either the significance or nature of a thing's role. And when, in this dissertation, I ask what role emotion plays in moral judgment, I mean to be asking *both* sorts of questions.

I shall first want to determine the *significance* of the role that emotion plays in moral judgment. To that end, in chapter III, I assess empirical evidence for thinking that emotion figures centrally in moral judgment. At the end of the day, I agree with many in thinking that emotion has a hugely significant role to play; but as we'll see, it will be important to review the evidence through the lens of the account of emotion defended in this chapter. When studies purport to show evidence of emotional stimulation

influencing moral judgment, for instance, we'll have to look closely at what is in fact being stimulated in these cases—maybe, in light of what I've argued in this chapter, it is indeed emotion; but maybe it is not. When it is claimed that people with emotional deficits of various sorts—e.g., psychopaths, people with autism, etc.—have consequent struggles related to moral judgment, we'll want to ask, in light of this chapter's discussion, what it is that grounds the assumption that the relevant psychological malfunction in these cases is indeed an *emotional* one. This is one reason it has been important to develop an account of what emotion is.

Once I've examined the evidence for thinking that emotion plays a very significant role in moral judgment, I'll move on in chapter IV to examine the exact *nature* of the role that it plays. In a paper recently submitted by one of my undergraduate students, the author writes, “[T]hese emotions anger and hate are, in the end, just emotions. Emotion should not factor into moral decision-making, as it clouds one's ability to make objective, rational decisions.” As I suggested in the last chapter, this is a very popular way of answering the question about the *nature* of the role that emotion plays in moral judgment and decision-making, and it is one that is by no means limited to non-philosophers. Walter Sinnott-Armstrong, for instance, has recently argued that the influence of emotion upon moral judgment often (if not always) renders moral judgment unreliable.⁴⁰ (He, too, uses the term ‘cloud’ to describe the sort of influence that emotion has upon moral judgment.)

If emotions were just feelings of changes in the body, then I could perhaps see why someone might think of them as distractions to clear (moral) thinking. But as I've

⁴⁰ Sinnott-Armstrong 2006.

argued in this chapter, this is not what emotions are. They involve construals and concerns, both of which are subject to rational and/or moral evaluation. So it seems there is at least some sense in which we can *get things right (or wrong)* when responding emotionally. And when things have gone right—when we are seeing things as they are, when our concerns are as they ought to be, etc.—why think that our emotions are *clouding* judgment? Why not think that emotions could contribute positively, in some epistemic sense, to moral judgment? These are the sorts of questions that I’ll explore in chapters III-V, and my answers will depend heavily upon the way I’ve conceived the nature of emotion here in chapter II.

III PSYCHOLOGICAL SENTIMENTALISM

At the end of the last chapter, I described two sorts of questions that one might ask about the role of emotion in moral judgment. The first has to do with *significance*: How significant is the role that emotion plays in moral judgment? Does emotion play a major role, a minor role, or no role at all? The second question has to do with the exact *nature* of the role that emotion plays in moral judgment: Is emotion a distraction to moral judgment? Does it cloud moral judgment? Or might emotion's contribution to moral judgment be something more positive? In this chapter, I will address the question of significance, and my answer will be that emotion has a very significant role to play. Indeed, emotion *figures centrally* in moral judgment.

I begin in §1 by explaining what I think it means to say that emotion “figures centrally” in moral judgment. As we will see, there are quite a number of ways in which one might say that emotion figures centrally in moral judgment, only one of which is of interest to me in this dissertation. On the view that I defend—sometimes called *psychological sentimentalism*—emotion figures centrally in moral judgment in that it is (among) the primary cause(s) of our moral intuitions. Then, in §§2-4, I discuss work in empirical moral psychology that I think amounts to a compelling case for psychological sentimentalism.¹ In §5, I critique what is arguably the most serious rival to the view that

¹ During my discussion of this empirical research, I will try to make note, whenever possible, of how my account of emotion from chapter II can help to explain what is going on in the relevant studies. As

I endorse—namely, a dual-process model of moral judgment. And finally, in the concluding section, I set the stage for a discussion of the epistemic status of moral intuitions, which shall be the focus of chapters IV and V.

Before getting into my discussion, however, I should issue a brief reminder about how I'll be using two key terms. As I explained at the end of chapter I, I use the term 'moral intuition' to refer to moral beliefs that are formed quickly and effortlessly, not on the basis of any kind of explicit inference. We do not *arrive at* our moral intuitions; rather, ordinarily, we simply *have them*—often despite paying little, if any, attention to our reasons for having them, or the process by which they were formed. In this way, I think our moral intuitions strongly resemble perceptual beliefs. I do not infer my way to the belief that there's a computer before me right now; rather, I simply have it—typically without paying any attention to my reasons for having it, or the process by which it was formed. (I'll revisit this resemblance between moral intuitions and perceptual beliefs in the next chapter.) And I use the term 'moral judgment' to refer to the process that is ordinarily (causally) responsible for the formation of moral intuitions, whatever that process might involve exactly. Again, I do not assume that moral intuitions are always formed in exactly the same way. But I do assume that there is something that can be identified as the process by which moral intuitions are ordinarily formed, and that is what I am calling 'moral judgment'. Starting with this chapter, it will be especially important to keep these definitions in mind.

far as I can tell, this is the best that can be done *at this point* to confirm that mine is an empirically viable account of the nature of emotion, and that these are indeed studies of emotion *as I conceive it*. But it will nonetheless be open to one to ask, of any one of these studies, "Why think that this is a study of emotion *as conceived in chapter II?*" And there may be very little that I could say in response. In my defense, I actually think it would be fair to ask this question with respect to *any* philosophical theory of emotion. And until philosophers and psychologists do more work *together* on this topic, the answers are likely to be less than satisfactory.

1. What ‘figures centrally’ means

The general claim that emotion figures centrally in moral judgment is an old one, and one that has been spun in a number of different ways. Very roughly speaking, it might be understood as a *semantic* claim (about the meanings of moral terms in thought and speech), a *metaphysical* claim (about moral properties), a *psychological* claim (about the causal grounds of moral intuitions), or an *epistemological* claim (about epistemic grounds for moral intuitions). In the hands of ethical expressivists, for instance, the idea that emotion figures centrally in moral judgment is one that has to do with the semantic content of the products of moral judgment, either in thought or in speech. Moral language, expressivists tell us, functions to express non-cognitive states like emotions.² To think an action wrong, then, is to have a thought the meaning of which is in some sense a matter of the emotion (or similarly non-cognitive state) expressed by the term ‘wrong’. This is the sense in which an expressivist might claim that emotion figures centrally in moral judgment.

Historically, the claim that emotion figures centrally in moral judgment has perhaps most often been understood as a claim about the status of moral properties, i.e., about that in which moral properties consist. Hume begins his *Enquiry Concerning the*

² I should note that there is an interesting discussion to be had about whether or not emotions are in fact *non-cognitive*. In the hands of feeling theorists, emotions sure seem to be non-cognitive states—no more cognitive, at least, than things like stomach aches and dizziness. But in what sense are evaluative judgments non-cognitive? In fact, judgment theorists like Solomon are sometimes called *cognitivists* about emotion. So one might consider the account of emotion that I defend in chapter II a kind of *hybrid* theory. I intentionally avoided using the labels ‘cognitive’ and ‘non-cognitive’ in chapter II, in part because I think it can lend more confusion than clarity to discussions of emotion theories. But it may nonetheless be worthwhile to ask whether emotions themselves ought to count as cognitive or non-cognitive states. However, I will not pursue this matter here.

Principles of Morals by remarking upon a controversy that he takes himself to have inherited

concerning the general foundation of MORALS; whether they be derived from REASON, or from SENTIMENT; whether we attain the knowledge of them by a chain of argument and induction, or by an immediate feeling and finer internal sense.³

It may sound as if Hume's interests in the role of emotion in moral judgment are primarily epistemological in nature—"whether we attain the *knowledge* of them by ... an immediate feeling and finer internal sense"—but by the end of the *Enquiry*, it is clear that Hume believes emotions (or sentiments) figure centrally in the make-up of moral properties.

The hypothesis which we embrace is plain. It maintains, that morality is determined by sentiment. It defines virtue to be *whatever mental action or quality gives to a spectator the pleasing sentiment of approbation*; and vice the contrary.⁴

So in Hume's hands, the claim that emotion figures centrally in moral judgment is ultimately a claim about the nature of morality itself. It's not just that *thinking* a thing virtuous involves responding to it emotionally in some way; rather, Hume is claiming, a thing's *being* virtuous depends upon the emotional responses that it engenders. Moral judgment depends upon the emotions because *morality* depends upon the emotions.

The view that morality depends upon the emotions is often called *philosophical sentimentalism*, and it is a view that is still quite popular today. Prinz, for instance, shares Hume's belief that moral properties are response-dependent properties. He writes, "When I say that something is wrong, I refer [...] to the property of causing emotions of

³ Hume 1983 [1751]: 13.

⁴ Hume 1983 [1751]: 85.

blame in me.”⁵ Other sentimentalist views hold that moral properties are dependent upon emotional responses in a slightly different way. According to so-called *neosentimentalist* views, things have the moral properties they have in virtue of its being *appropriate* in some sense (e.g., rational) to have certain emotions in response to them. So for instance, Allan Gibbard writes, “What a person does is *morally wrong* if and only if it is rational for him to feel guilty for doing it, and for others to resent him for doing it.”⁶ What all of these views have in common is the idea that emotion figures centrally in moral judgment in the sense that the *domain* of moral judgment—morality itself—depends upon the emotions.

Compared to ethical expressivism and philosophical sentimentalism, the view that I wish to defend in this chapter is modest. In fact, when it comes to the semantics of moral claims and the metaphysics of moral properties, the view to be defended here is entirely noncommittal. In the remainder of this chapter, I’ll defend a version of what is sometimes called *psychological sentimentalism*—the view that the process by which our moral beliefs are ordinarily formed is one in which the emotions figure centrally.

It may be helpful, in order to really draw out the contrast between psychological sentimentalism and the other two views, to consider an analogy to perceptual judgment. Suppose I grab a pot off the stove to clean it, unaware that it was recently used to boil water and is still very hot. The pot burns my hand, and I shout (or just think to myself), “That’s hot!” According to psychological sentimentalism, the role that emotion plays in moral judgment is very much like the role that perception—in this case, the burning

⁵ Prinz 2006: 35.

⁶ Gibbard 1990: 42.

sensation—plays in my judgment of the temperature of the pot. The burn has nothing whatsoever to do with the *semantic content* of the claim “That’s hot!” Nor does the relevant aspect of reality—i.e., the temperature of the pot—*depend* upon the burn in the way the Humean thinks morality depends upon the sentiments. But the burn does *cause* me to judge that the pot is hot—I judge that it’s hot *because* of the burn. Likewise, the psychological sentimentalist says, in most ordinary cases at least, we judge things morally as we do *because* of our emotional responses to them.

So for my purposes, the sense in which emotion “figures centrally” in moral judgment is this: in ordinary cases of moral judgment, the *products* of moral judgment—moral intuitions—are the *effects* of emotion, i.e., we have the moral intuitions that we have *because* of our emotional responses to things. As I’ve explained, this is a semantically and metaphysically innocent claim, and I take that innocence to be an advantage that psychological sentimentalism has over the two other views mentioned. But psychological sentimentalism does have interesting *epistemological* implications: if moral intuitions are ordinarily the causal consequences of our emotions, then what is the *epistemic status* of these intuitions? Assuming my physiology is in healthy, working order, presumably, the burning sensation I feel upon touching the pot contributes positively to the justification of my immediate and unreflective (i.e., intuition-like) belief about its hotness. But can the same be said of emotions and their role in moral judgment? I’ll answer this question in the next chapter.

In what remains of this chapter, however, I will address the following question: What evidence is there for thinking that emotion figures centrally in moral judgment in the way that psychological sentimentalism claims that it does?

2. Moral dumbfounding

It may help to step back a bit and recall some of the history behind this debate. For the better part of the 20th century, the most influential approach to moral judgment was the one endorsed by the cognitive developmental tradition. Proponents of this tradition—most notably Jean Piaget and Lawrence Kohlberg—generally affirmed some version of *psychological rationalism*, which is the view that moral intuitions—or moral beliefs, more generally⁷—are ordinarily the products of a conscious reasoning process, often involving the application of general moral principles to particular cases. As these theorists describe it, moral judgment is a kind of *skill* that develops through a number of stages. In the beginning, young children judge things morally in accordance with a set of moral rules, typically taught and enforced by a parent, or some other authority figure. But by the time people reach full moral maturity (stage 6, according to Kohlberg), they have internalized principles of a just society, and are able to apply those principles to particular actions, and even to the rules and laws that govern society. Kohlberg conceived of the fully developed moral judge as essentially the realization of a Kantian or Rawlsian ideal—someone who consciously thinks in terms of such things as respect for persons, equal rights, and universalized principles of action.

⁷ Note that, as I define the term in chapter I, it is apparently impossible for moral *intuitions* to be the products of a conscious reasoning process. As I see things, this is indicative of a flaw in the cognitive developmental approach to moral judgment, and not in my definition of the term ‘moral intuition’. I say more about this below, but the gist is that I believe that cognitive developmental theorists like Piaget and Kohlberg were actually investigating something *other* than moral judgment (and moral intuitions) all along.

This general approach to moral judgment has since been attacked from quite a number of angles. For my part, I want to focus on the assumption made by psychological rationalists that moral judgment is ordinarily—or even *often*, for that matter—a process in which conscious reasoning figures centrally. Kohlberg’s primary source of evidence for his theory, by far, comes from an impressive battery of moral judgment *surveys*, administered to subjects in studies that sometimes spanned more than two decades. Basically, Kohlberg and his colleagues just assumed that *whatever considerations subjects cited in justifying their responses on these surveys were the very considerations about which subjects consciously reasoned when formulating their responses*. But in the decades since the completion of Kohlberg’s most influential study, social psychologists have found strong evidence for thinking that there is not this sort of connection between considerations cited by a believer in order to justify belief *b* and the process by which *b* was originally formed.

In 1978, for instance, psychologists Timothy Wilson and Ross Nisbett published the results of a now famous study in which subjects were asked to choose from among 4 pairs of nylon pantyhose, and then explain or justify their preference. However, unbeknownst to the subjects, all 4 pairs of pantyhose were actually manufactured by the same company, and indeed pulled from the same package. Despite expressing sometimes very firm beliefs in the superior quality of one pair over the others, the subjects’ judgment of pantyhose quality in this case was clearly influenced to a large degree by the *physical position* of the pantyhose. Of the 4 identical pairs, 12% of participants judged the left-most pair to be best, 17% preferred left-center pair, 31% preferred the right-

center pair, and 40% preferred the right-most pair. While subjects offered no fewer than 80 reasons in total for their preferences, not a single one of them mentioned the position of the pantyhose as a factor contributing to their judgment. And in fact, when asked if such a thing could have influenced their judgment of the pantyhose, 51 of the 52 subjects insisted that it could not.⁸

On the basis of studies like this one—and there are plenty of them—psychologists in general are wary of drawing conclusions about the causal genesis of people’s beliefs from the sorts of considerations that people cite when asked to justify their beliefs. Until recently, however, these concerns hadn’t been explored in conjunction with *moral* judgment. Is there any evidence for thinking that moral intuitions are often the causal effects of things quite different from the sorts of considerations that people cite when asked to justify those intuitions? If so, this would apparently be a serious problem for psychological rationalists like Piaget and Kohlberg.

For a series of studies, psychologist Jonathan Haidt and colleagues concocted a number of hypothetical scenarios meant to describe actions that people will find offensive, but harmless.⁹ In one scenario, for instance, a woman cuts her national flag (e.g., the American flag) into pieces and uses the rags to clean her toilet. In another, a family thoroughly cooks and eats a pet dog after a car strikes and kills the dog in front of their house. After reading the vignettes, subjects were asked six probe questions,

⁸ Wilson and Nisbett 1978: 123-124. The one subject who admitted that position could have influenced her evaluation also noted that she was currently taking three psychology courses.

⁹ It may be worth noting that Haidt’s primary interest with these studies had more to do with the boundaries of the moral domain than the psychology of moral judgment—e.g., we generally treat *harmful* actions as morally significant, but what about *harmless* actions that are nonetheless *disgusting*?

including: (a) *Evaluation*: “What do you think about this? Is it very wrong, a little wrong, or is it perfectly OK for [act specified]?” (b) *Justification*: “Can you tell me why?” (c) *Harm*: “Is anyone hurt by what [the actor] did? Who? How?” Here is how Haidt describes what turned out to be a common occurrence:

Participants often stated immediately and emphatically that the action was wrong, and then began searching for plausible reasons. Participants frequently tried to introduce an element of harm, for example by stating that eating dog meat would make a person sick, or by stating that a person would feel guilty after voluntarily using her flag as a rag. When the interviewer repeated the facts of the story (e.g., that the dog was thoroughly cooked so no germs were present), participants would often drop one argument and begin searching for another.¹⁰

So, rather than abandon their initial, intuitive response to the case, subjects would just search for other reasons to support it. And whenever the search *failed*, subjects typically still would not give up their initial intuition about the case, instead offering what Haidt calls a “statement of moral dumbfounding,” i.e., a “statement to the effect that they thought an action was wrong but they could not find the words to explain themselves.”¹¹ For another study, Haidt and Hersh asked conservatives and liberals a number of questions about issues related to sexual morality—e.g., questions about homosexuality, incest, and unusual forms of masturbation. Here again, subjects were often morally dumbfounded: “[T]hey would stutter, laugh, and express surprise at their inability to find supporting reasons [for their moral intuitions], yet they would not change their initial judgments of condemnation.”¹²

¹⁰ Haidt, Bjorklund, and Murphy unpublished: 3.

¹¹ Haidt, Bjorklund, and Murphy unpublished: 9.

¹² Haidt 2001: 817.

How should we interpret the results of these studies? According to Haidt, one thing that these studies show is that, while conscious reasoning may have *something* to do with moral judgment, it does *not* play the sort of role that psychological rationalists like Piaget and Kohlberg assume. Compare two stereotypes, one of a scientist and the other of a lawyer. The scientist collects data and then forms whatever conclusion is best supported by the data. The lawyer, on the other hand, does the opposite: he begins with a conclusion—e.g., that his client is innocent—and then searches for evidence for this conclusion. As Haidt sees things, his studies—and in particular, cases of moral dumbfounding—strongly suggest that the role that conscious reasoning plays in moral judgment is more like that of the lawyer than the scientist. That is, people arrive at moral intuitions as the result of unconscious mental processes; it is only when they are asked to justify the intuitions that conscious reasoning steps in to serve as a kind of defense attorney, searching for reasons that support the original intuition. The only difference between cases in which a person is morally dumbfounded and cases in which a person is not, then, is just that, in the latter cases, conscious reasoning is *successful* in its search for supporting reasons.

If Haidt's interpretation is on point, then the psychological rationalist is just wrong: moral judgment is not a process in which conscious reasoning figures centrally. Rather, conscious reasoning only enters the picture *after the intuition has already been formed*. Indeed, it is hard to imagine how someone like Kohlberg would interpret the results of Haidt's studies. How could subjects be unable to provide reasons for beliefs in which they are so confident, if the beliefs themselves are supposed to be the products of a process of conscious reasoning? Are these people reasoning their way to moral

intuitions, and then immediately forgetting how it was that they arrived at their newfound beliefs? And why are they still so confident in their moral intuitions, even when, *by their own admission*, they cannot provide any reasons in support of those beliefs?

Now, Haidt's moral judgment studies do not by themselves provide particularly strong support for psychological sentimentalism. It is one thing to problematize the rationalist's claim that conscious reasoning figures centrally in moral judgment, as I believe these studies do. It would be quite another thing, however, to show that *emotion* figures centrally in moral judgment. After all, it would be perfectly consistent with the results of Haidt's studies to think that moral judgment ordinarily involves a non-emotional psychological process that just happens to operate at a subconscious level.¹³ If people's moral intuitions are ordinarily the products of a mental process that is entirely *non-emotional*, but also unconscious, then it should not come as a surprise that subjects are often unaware of what causal factors first led them to have the beliefs that they have. So, even if we grant that moral intuitions are ordinarily the products of an *unconscious* process, what evidence is there for thinking that it is an *emotional* one?

In the next two sections, I'll discuss studies that more directly support the sentimentalist's claim that it is *emotion* that figures centrally in moral judgment.

¹³ Marc Hauser, Susan Dwyer, and Bryce Huebner defend such an account of moral judgment (see, e.g., Hauser 2006; Dwyer 2009; and Dwyer, Huebner, and Hauser 2010). They call it the "Linguistic Analogy," since it builds upon an analogy to Chomsky's suggestion that underlying the development of our capacity for language is a *universal grammar*, i.e., an innate set of principles of grammaticality, and the like. Likewise, Hauser, Dwyer, and Huebner argue, moral judgment is guided by a kind of universal moral grammar, i.e., an innate set of moral principles against which we unconsciously analyze actions and events in order to arrive at moral intuitions. Among the evidence they cite in defense of the Linguistic Analogy is the phenomenon of moral dumbfounding.

3. Evidence of the effects of emotional stimulation upon moral judgment

Suppose a chemist is trying to determine what sort of role, if any, substance A plays in chemical reaction X. She might do a number of things to test this: she might increase or decrease the level or amount of A, to see if this has any effects upon X; she might introduce A to environments in which X would not otherwise occur, to see if A causes X there as well; or she might remove A entirely, to see if X continues, unaffected, in the absence of A. Other things being equal, if changes in the amount of A result in corresponding changes to the frequency, intensity, or some other aspect of X; if the introduction of A to new environments results in the occurrence of X there as well; and if removing A causes X to change drastically, or to cease altogether—this would all amount to compelling evidence for thinking that A figures centrally (as a cause, or catalyst) in the production of X.

Interestingly, psychologists and neuroscientists have come up with ways of simulating these sorts of tests in order to determine how significant a role, if any, emotion plays in moral judgment. In this section, I'll discuss studies in which the levels or intensity of subjects' emotions are adjusted in various ways, with notable effects upon moral judgment. And I'll also discuss a study in which subjects' emotions are stimulated in a context in which neither such emotions nor moral judgment would typically occur. Then, in the next section, I'll discuss studies involving people with an emotional deficit of some sort.

There are a number of recent studies in which an artificial induction, or increase, in some negative emotion has been shown to result in a corresponding increase in the

harshness or severity of people's moral judgment, i.e., an increase in the degree to which they condemn a particular action or actor. In one such study, for instance, subjects were asked to complete a moral judgment survey in which they were presented with six vignettes—three involving a disgusting violation (a man eating his dead dog, plane crash survivors considering cannibalism, and a man deriving sexual pleasure from playing with a kitten), and three involving violations unlikely to induce disgust (e.g., a man intentionally falsifying his resume). Some of the subjects were directed to complete the survey at a particularly tidy and clean workspace, while others were directed to complete the survey at a workspace likely to induce a disgust response:

An old chair with a torn and dirty cushion was placed in front of a desk that had various stains and was sticky. On the desk there was a transparent plastic cup with the dried up remnants of a smoothie and a pen that was chewed up. Next to the desk was a trashcan overflowing with garbage including greasy pizza boxes and dirty-looking tissues.¹⁴

Experimenters predicted that subjects seated at the dirty desk would judge the actors described in the vignettes more harshly than subjects seated at the clean desk, and this is precisely the effect that was observed—especially when the experimenters factored in subjects' attentiveness to their own internal states, as measured on a Private Body Consciousness (PBC) scale.

Planned comparisons showed that when averaging across all six vignettes, high-PBC participants in the disgust condition perceived the actions of the story character as more wrong than participants in the no-disgust condition. [...] In contrast, for low-PBC participants, ratings of moral condemnation did not differ between the disgust and no-disgust conditions.¹⁵

¹⁴ Schnall et al. 2008: 1101.

¹⁵ Schnall et al. 2008: 1101.

In another study, the same experimenters asked subjects to complete a moral judgment survey in an outdoor setting. This time, some subjects were seated approximately 6 feet away from a trash container in which a particularly odorous substance had just been sprayed.¹⁶ Here, too, subjects seated near the disgusting smell judged the actions, or actors, described in the vignettes more harshly than did subjects seated in an area with no such odor.¹⁷

Another method used to artificially induce disgust responses in subjects is that of posthypnotic suggestion.¹⁸ In a study conducted by Haidt and Thalia Wheatley, sixty-four highly hypnotizable subjects took part in a group hypnosis session that included a posthypnotic suggestion to feel “a brief pang of disgust” upon reading the word ‘often’, but to have no memory of this instruction until being prompted to remember (after

¹⁶ The substance used was ammonium sulfide in a water solution. Ammonium sulfide may be best known for its presence in such popular prank items as “stink bombs” and “fart spray.”

¹⁷ Schnall et al. 2008: 1097-1099.

¹⁸ While it may be worth noting that hypnosis is a notoriously controversial practice, I think there are few who doubt that it can be an effective means of altering such things as a person’s self-image, experience of pain, and disposition to affective states like emotions. How does all of this happen, exactly? This is a matter of some controversy. The picture painted by the most widely accepted theory of hypnosis in the clinical domain is quite complex, involving the manipulation by the hypnotist of divided and dissociated levels of consciousness (see, for instance, Alladin 2008: ch. 1). In light of my discussion in chapter II, though, it may be helpful to understand hypnotic suggestion in terms of a subconscious conditioning of a patient’s *construals*. In fact, hypnotherapists regularly speak of using hypnosis in order to alter *the way a patient sees something*. One popular form of hypnotic suggestion, for instance, is known as “ego-strengthening,” and it is characterized as a means of helping patients “view themselves as self-efficacious ... in order to develop feelings of confidence and optimism and an improved self-image” (Alladin 2008: 45). Hypnosis can be used to manipulate people’s emotions, then, because it is an effective means of altering their *construals* of various things (e.g., themselves). If this is right, then we might interpret Wheatley and Haidt’s study along something like the following lines. Rather than feeling “a brief pang of disgust” upon reading some word and then *using* that disgust response “as information for moral judgment,” as the experimenters themselves put it, what is actually going on is that subjects are *seeing* things differently (e.g., seeing the congressman as even more corrupt than they would have otherwise seen him), and then simply *judging things as they see them*.

completion of the survey).¹⁹ Subjects were then given a series of vignettes describing some sort of moral transgression, and asked to rate both “how morally wrong” and “how disgusting” they considered the action described in the vignette “by making a slash mark along a 14-cm line anchored by the endpoints *not at all morally wrong* and *extremely morally wrong* or *not at all disgusting* and *extremely disgusting*.”²⁰ The slash marks were later converted to number on a scale from 0 to 100. Importantly, the subjects were unaware of any connection between the hypnosis session and the moral judgment task.

The results suggest that even hypnotically induced disgust can cause people to judge actions more harshly than they would otherwise judge them. Here, for instance, is one of the vignettes:

Congressman Arnold Paxton frequently gives speeches condemning corruption and arguing for campaign finance reform. But he is just trying to cover up the fact that he himself [will take bribes from/is *often* bribed by] the tobacco lobby, and other special interests, to promote their legislation.

Some subjects read the vignette with the hypnotic cue word ‘often’, while other subjects read the vignette without the cue. When asked how *disgusting* the congressman’s behavior is, those who read the vignette *without* the cue word gave it an average rating of 38.92, while those who read the vignette *with* the cue rated it much higher, at 72.37. When asked how *morally wrong* the congressman’s behavior is, those who read the vignette *without* the cue word gave it an average rating of 78.73, while those who read the vignette *with* the cue gave it an average rating of 91.28. In response to a case describing a sexual encounter between second cousins, those who read the vignette

¹⁹ Wheatley and Haidt 2005. Of the 64 participants, 19 remembered the posthypnotic suggestion prior to being prompted to do so. Wheatley and Haidt limited their analysis to the remaining 45 subjects.

²⁰ Wheatley and Haidt 2005: 780.

without the cue give it a disgust rating of 43.24, while those who read the vignette with the cue rated it much higher at 72.46. And again, the presence of the disgust cue increased ratings of moral wrongness, this time from 43.29 to 67.63. So the posthypnotic suggestion seems to have succeeded in heightening subjects' disgust response, and the heightened disgust response appears to have made people significantly harsher as moral judges.

One should not, however, conclude that there is something especially efficacious about *disgust* in this regard. In a study by Lerner and colleagues, subjects were asked to read 4 counterbalanced vignettes, all of which described some kind of harm resulting from worker negligence.²¹ Then, after reading the vignettes, subjects completed a series of Likert 7-point scales assessing their reactions to each case, including 5 items specifically designed to assess subjects' judgment about the degree to which the workers are responsible and ought to be punished. Just prior to reading the vignettes, however, some subjects were shown a video clip likely to induce an *anger* response (featuring a bully humiliating and beating up a teenager), while others were shown an emotionally neutral video clip (featuring abstract shapes and colors). Here, too, subjects whose anger response had just been elicited by the video judged the workers more harshly than did subjects who watched the emotionally neutral video. Psychologists have also found that inducing *sadness* in subjects can cause them to evaluate others more harshly than they otherwise would have.²²

²¹ Lerner et al. 1998.

²² Fogas and Bower 1987.

Nor should one conclude that it is only the stimulation of *negative* emotional states that can have effects upon people's moral judgment of others. Whereas inducing anger can lead people to be harsher in their judgment of the degree to which another ought to be punished, inducing a more positive affective response can apparently have precisely the opposite effect. We know that physical attraction is likely to induce positive affect, and there are several studies according to which the more physically attractive we find a person, the more lenient we will be in our estimation of his or her deservingness of punishment.²³ And while sadness can apparently lead to more negative appraisals of others, some studies suggest that the positive affective response that accompanies physical attraction can lead to more positive appraisals of others, including attributions of such virtuous traits as honesty and integrity.²⁴ Psychologists can also induce positive emotional states (or reduce the intensity of negative emotional states) by showing subjects a humorous video clip; and in one study, subjects who viewed the humorous clip were significantly *less* harsh in their responses on a moral judgment survey than subjects who viewed an emotionally neutral video clip.²⁵

Many of the studies that I've just described involve an increase in levels of emotion that likely already would have been present. Even in the absence of the hypnotic cue word, for instance, subjects still were disgusted by the congressman and believed that he behaved wrongly. So the evidence may so far only provide a small amount of support to psychological sentimentalism—it *does* show that emotions can causally influence

²³ See, e.g., Darby and Jeffers 1988.

²⁴ Dion et al. 1972.

²⁵ Valdesolo and DeSteno 2006. The humorous clip was from the comedy sketch program "Saturday Night Live," and the emotionally neutral clip was taken from a documentary about a small Spanish village.

moral judgment, but, strictly speaking, anyone can admit that our emotions may sometimes cause us to judge things *more harshly*. What would *really* be interesting—and would provide even stronger support for sentimentalism—would be if there were evidence of emotional stimulation causing people to judge things morally when they otherwise would not judge them morally *at all*.

In a follow-up to their first experiment with hypnotic disgust, Wheatley and Haidt again induced a disgust response to the word ‘often’ via posthypnotic suggestion; then had subjects read and respond to vignettes describing various sorts of actions; and again asked subjects to rate “how morally wrong” and “how disgusting” they considered the actions described in the vignettes. This time, however, they included a vignette in which someone performs an action of no moral significance whatsoever:

Dan is a student council representative at his school. This semester he is in charge of scheduling discussions about academic issues. He [tries to take/*often* picks] topics that appeal to both professors and students in order to stimulate discussion.

When asked how *disgusting* they consider Dan’s behavior, those who read the vignette *without* the cue word gave it an average rating of 2.3; and when they were asked how *morally wrong* they consider Dan’s behavior, the same subjects gave it an average rating of 2.7. (Again, both of these numbers are on a scale of 0 to 100. And given the researchers’ somewhat imprecise method for collecting, and then assigning numbers to, subjects’ ratings, a rating of 2.7 might as well be a rating of 0.) On the other hand, when asked how *disgusting* they consider Dan’s behavior, subjects who read the vignette *with*

the cue rated it much higher, at an astonishing 20.9;²⁶ and when they were asked how *morally wrong* they considered Dan's behavior, subjects gave it an average rating of 14.0.²⁷ So, more than just causing people to morally judge things *differently* (e.g., more harshly) than they otherwise would, it appears as if emotional stimulation can even cause people to judge things morally *when they otherwise would not judge them morally at all*.

Apparently, changes in the intensity of certain emotions can result in corresponding changes to the intensity of people's moral judgment. Furthermore, the stimulation of certain emotions can even cause people to engage in moral judgment when, presumably, they otherwise would not. Here, too, this alone is not sufficient to make the case for psychological sentimentalism. But I do think a picture is beginning to emerge of the (largely) unconscious psychological processes that ordinarily underlie the production of moral intuitions, and it is one in which emotion appears to play a significant role. To return to the analogy at the beginning of the section, suppose the chemist finds that changes to the amount of substance A lead to corresponding changes in reaction X, and also that the introduction of A to other environment causes X to take place there as well. If this is some evidence for thinking that A figures centrally as a cause of X, then, in this section, we've likewise seen some evidence for thinking that emotion figures centrally as a cause in the production of moral intuitions.

²⁶ For comparison, in the first study, subjects who read a vignette involving shoplifting (without the cue word) gave it an average disgust rating of just 19.79. So these subjects apparently find Dan's (totally morally neutral) behavior every bit as disgusting subjects in the control group find shop-lifting.

²⁷ In discussion with the researchers, one participant who read the vignette *with* the hypnotic cue described Dan as a "popularity-seeking snob," and another said, "It just seems like he's up to something."

4. Evidence of the effect of emotional deficit upon moral judgment

Here is another useful analogy. During the 2009 and 2010 football seasons, the Indianapolis Colts won 24 of 32 regular season games, had more total passing yards than any other team in the NFL, made it to the playoffs each season, and even to the Super Bowl at the conclusion of the 2009 season. But after the 2010 season, Colts quarterback Peyton Manning had surgery on his neck, and was unable to play a single game of the following season. In 2011, the Colts won only 2 of 16 games, finished near the very bottom of the NFL in passing yards, and of course, did not make it to the playoffs. What sort of role did Peyton Manning play in the Colts' success? Apparently, a *very* significant one. By the conclusion of the 2011 season, in light of the Colts' dismal performance, sports commentators and journalists joked that Manning should be named the NFL's Most Valuable Player for that season, despite having not played in a single game. Sometimes, it seems, the best way to gauge the causal significance of x to y may simply be to remove x , and see what happens to y .

How might this lesson be applied to the subject at hand—i.e., the causal significance of emotion to moral judgment? If it turns out that an abnormal *absence* of emotion results in similarly abnormal effects upon moral judgment, this would lend further support to the sentimentalist's claim that ordinary moral judgment is a process in which emotion figures centrally (just as their performance without Manning in 2011 might lend support to the claim that he "figured centrally" in the Indianapolis Colts' success in prior seasons). In order to see the effects of "an abnormal absence of emotion," we should look to people with conditions characterized by some sort of emotional deficit—i.e., a systematic absence or blunting of normal emotions.

Psychopathy is a type of anti-social personality disorder characterized by, among other things, a lack of guilt, remorse, and empathy. What underlies this emotional deficit? According to Blair, the psychopath's lack of guilt, remorse, and empathy likely has something to do with a lack of concern for the wellbeing of others.²⁸ This comes as no surprise, in light of my discussion from chapter 2: underlying the psychologically normal person's empathetic emotional response, for instance, is both (a) a construal of another as being in some kind of distress, and (b) a concern for the wellbeing of the other (to which the construal relates disharmoniously). Even if psychopaths are capable of (a)—and I assume they are—if they are unconcerned with the distress of others (or at least, much *less* concerned than is the normal person), they will consequently be unmoved emotionally by the sight of another in distress.²⁹

Does this emotional deficit affect a psychopath's capacity for moral judgment in any way? At first glance, it appears not. After all, psychopaths are apparently just as likely to call something “wrong” as are non-psychopaths. Consider, for instance, the following quote from serial killer Ted Bundy:

It was almost as if he [I] said it was wrong for all these things to happen. “It is wrong for me to jaywalk. It is wrong to rob a bank. It is wrong to break into other people's houses. It is wrong for me to drive without a

²⁸ Blair 2007.

²⁹ This example may put more pressure on Roberts (see ch. II, n. 35, and the relevant portion of the text). If concerns must be built into construals in the way that Roberts thinks, then he apparently cannot allow that psychopaths see a person as distressed, in the way that non-psychopaths do. My account of emotion, however, can easily allow for this, explaining the emotional differences entirely in terms of a difference of *concern*.

driver's license. It is wrong not to pay your parking tickets. It is wrong not to vote in elections. It is wrong to intentionally embarrass people.”³⁰

Bundy *says* all the right things—it *is* wrong to rob banks, it *is* wrong to break into other people's homes, etc.—and so, it can be easy to assume that, regardless of any emotional or motivational deficits that Bundy may exhibit, his capacity for moral judgment is intact. But in this case, appearances are deceiving.

Imagine someone who is totally colorblind, but nonetheless able to detect different wavelengths of light (perhaps by means of an instrument he carries with him). So while he cannot see red, he can know when light has a wavelength of about 650nm (the wavelength of visible red light). Further, imagine that he learns that others use the term ‘red’ to describe things that radiate or reflect light with a wavelength of 650nm—and so forth for the other color terms. Just as Bundy seemed to *say* all the right things *morally*, our colorblind companion is likely to *say* all the right things when it comes to judging the colors of things. But the process out of which his color beliefs are formed—i.e., his color *judgment*—is fundamentally different than that of the (perceptually) normal person. His color beliefs are, in a sense, parasitic upon the beliefs and practices of others—e.g., when he says, “That’s red,” it’s as if he is saying, “That’s what people call ‘red’.” Furthermore, this difference between his color judgment and that of the (perceptually) normal person will underlie various other failures of his associated with judgment about the colors of things—e.g., he’ll not understand the difference between warm and cool colors as we do; he’ll not understand what it means for colors to clash or coordinate; etc. So even if we were unaware of the process by which he forms his color

³⁰ Michaud and Aynesworth 1989: 116. For the record, at times during the interviews, Bundy would refer to himself in the third person. So while it may sound like he’s reporting the opinions of someone else, he is actually reporting his own thoughts here.

beliefs—i.e., unaware of his use of the wavelength-detection instrument—we'd be right to suspect that something about the process is abnormal, *defective* even, when he consistently attaches the labels 'warm' and 'cool' to colors incorrectly, when he consistently fails to match and coordinate colors, etc.

I think something very similar is going on with the psychopath's moral judgment. Psychopaths do appear to have moral beliefs, but the process out of which those beliefs are formed is apparently quite different from that of the psychologically normal person, and probably similarly parasitic upon the beliefs and practices of others. Intelligent, articulate, and observant as they are, psychopaths are capable of recognizing some of the norms that are in place around them—moral or otherwise—and also the sorts of linguistic constructions that people use when referring to violations of those norms (e.g., "That's wrong," "He shouldn't have done that," etc.). So just as my colorblind friend can look at a red light and correctly judge that it is red, a psychopath can witness someone intentionally embarrassing another and correctly judge that the person thereby behaves wrongly. But here again, even though the psychopath often gets things *extensionally* correct, there may still be reason to suspect that his capacity for moral judgment is deficient by comparison to the psychologically normal person.

One such reason, for instance, is the apparent failure on the part of psychopaths to distinguish between moral and conventional wrongdoing. Now, I think it would be a mistake to think that there is some sort of very clear line between these two categories of wrongs. But *in general*, it seems, moral transgressions are more serious, less permissible, and less likely to be contingent upon authority—e.g., hitting a fellow student is more

serious and less permissible than chewing gum in class; and even if the teacher explicitly allows both, it will still be wrong to hit. And from a very young age (around the third birthday), psychologically normal children seem to recognize the differences between moral and conventional wrongs along all three of these dimensions.³¹ In a study comparing psychopathic and non-psychopathic criminals, however, Blair found that psychopaths consistently fail to draw a significant distinction between moral and conventional wrongs along any of the three dimensions.³² Apparently, as far as the psychopath can tell, burping loudly in public is wrong *in the same way* that violent attacks are wrong.³³ And of course, another reason for thinking that psychopaths' capacity for moral judgment is deficient is their apparent failure to be motivated to act in accordance with their professed moral beliefs. Regardless of how close one happens to think the connection is between moral judgment and motivation, it is certainly natural to think that, if a person *really judges x* to be morally wrong, then we can expect her to be

³¹ See Smetana 1993.

³² Blair 1995.

³³ Interestingly, contrary to what Blair predicted, rather than treating moral transgressions as conventional, psychopaths actually treat conventional transgressions as moral (i.e., as relatively more serious, less permissible, and not contingent upon authority). This is especially interesting when one considers the sorts of justifications provided by the psychopathic and non-psychopathic subjects for their responses. Non-psychopathic subjects typically cited *harm to the victim* as among their reasons for judging an action along characteristically moral lines. Psychopaths, on the other hand, rarely made reference to the victim's welfare; instead, they made reference (whether implicitly or explicitly) to social rules, expectations, and the like (e.g., "It's not acceptable to do that"). So why do psychopaths judge most or all transgressions along characteristically moral lines, if they apparently do not have the same reasons for doing so as non-psychopaths? For an answer, Blair reminds us that the participants in his study were criminals serving time in prison: "These subjects were all incarcerated and presumably motivated to be released. All wished to demonstrate that the treatments they were receiving were effective. They therefore would be motivated to show how they had learned the rules of society ... [The psychopaths' inability to recognize a difference between moral and conventional transgressions], coupled with a desire to demonstrate an adherence to societal rules, results in their judgment of all the transgressions as *authority independent*. The non-psychopaths, in contrast, though presumably equally motivated to be released, are incapable of ignoring the distinguishing features of moral and conventional transgressions ... and thus answer the *authority jurisdiction* question appropriately" (1995: 23-24).

motivated to avoid doing *x*. But these expectations are consistently flouted by psychopaths, who appear to be prone to immoral (and often criminal) behavior.

So again, psychopaths may say the right things—they call moral and conventional transgressions “wrong,” just as we do—but the fact that they apparently do not see a moral difference between, say, breaking into people’s homes and driving without a driver’s license, coupled with their infamous motivational failures, suggests that there is a defect of some sort in their capacity for moral judgment. When psychopaths say that a thing is morally wrong, they are not giving voice to a moral intuition in the way that *we* are when we say that things are morally wrong. Rather, what the psychopath is doing is more akin to parroting the moral claims of those around him. And what underlies this defect in the psychopath’s capacity for moral judgment, many think, is the emotional deficit noted earlier.

Why think that it is an *emotional* deficit that underlies their deficient capacity for moral judgment? Well, for one thing, all of the psychopath’s more stereotypically “cognitive” capacities appear for the most part to be intact.³⁴ As I noted earlier, psychopaths are often extremely intelligent, articulate, and observant people. And for another thing, researchers have discovered similar moral judgment-related deficiencies in

³⁴ It might be argued that the psychopath’s apparent failure to recognize what is *moral* about a moral transgression has something to do with a deficient capacity for perspective-taking—i.e., the psychopath cannot “see” what makes intentionally embarrassing someone *morally* wrong because he cannot take the perspective of the person being wronged. And perspective-taking is often thought to be a *cognitive* (opposed to *affective*, or *emotional*) mental exercise. There are two ways for the psychological sentimentalist to say here. First of all, it is unclear that psychopaths really cannot take the perspective of another. Blair, for instance, believes that they can (see Blair et al. 1996), and that this explains why psychopaths are often so successful at manipulating others. And second, it is unclear that perspective-taking is really a *cognitive* mental exercise to begin with. Peter Hobson, for instance, makes a compelling case for thinking that the autistic child’s struggle with perspective-taking is rooted in a kind of *emotional* impairment (see Hobson 2002 and 2005).

patients with significantly different conditions—e.g., people with autism, people with injuries to particular parts of the brain, etc.—and what is common to all of these cases is that the patients exhibit an emotional deficit or impairment of some sort.³⁵

Unlike psychopaths, people with autism have a particularly difficult time recognizing the mental states of others.³⁶ In one study, for instance, two groups of children and adolescents—one with autism, and one without autism—were asked to watch a brief sequence of images in which points of light in the shape of people perform some action (gestures of surprise, sadness, fear, anger, and happiness), and then asked to identify “what’s happening” in the images.³⁷ Despite the fact that both groups of children correctly identified the point-light displays as *people*, all but one of the non-autistic children commented on the person’s emotions (surprise, sadness, etc.), while 10 of the 13 autistic children focused *only* on the person’s physical movements, never even mentioning the person’s emotional states.³⁸

In the case of the children and adolescents with autism, it was the person’s movements and actions rather than feelings that were reported. For example, they described the sad figure as ‘walking and sitting down on a

³⁵ For more on the moral judgment of psychopaths, see Nichols 2004: 11-20 and Prinz 2007: 42-47. For studies of the deficient moral judgment of patients with damage to their ventromedial prefrontal cortex, see Young et al. 2010. Such patients also exhibit motivational deficiencies similar to those of the psychopath (see Doris and Stich 2005: 123-128; see also Damasio’s description of the famous case of Phineas Gage in 1994: 3-33). And for studies of the deficient moral judgment of people with autism, see Grant et al. 2005 and Takeda et al. 2007). I go on to say a bit about the connection between autism and moral judgment in the text.

³⁶ See n. 34.

³⁷ Hobson explains why he and his colleagues elected to use point-light displays instead of, for instance, pictures or video of real human expressions: “In previous studies, we had found evidence that, at least on occasion, children with autism appeared to be able to name bodily expressed emotions, for example in faces, but by employing unusual perceptual/cognitive strategies that might not reflect full experience and understanding of what the expressions signify for a person’s feelings” (2005: 191).

³⁸ Hobson 2005: 192.

chair', 'walking and flapping arms and bent down', and 'walking and waving his arms and kneeling down [...] hands to face'.³⁹

Among the more consistent symptoms of autism, as is exhibited in these (and other) studies, is a deficient capacity to read the subjective states of other people off of their actions. And what is it that lies at the (causal) root of this deficiency? Many think that it stems from a socio-emotional impairment:

[C]hildren with autism are unusual in their relative failure to become emotionally engaged with the emotional states of other people. [...] Put simply, this limitation in the children's experience of intersubjective contact and reciprocal engagement with others is a fundamental impediment to growth in their understanding of what it is to have a subjective mental orientation.⁴⁰

In other words, due (in part) to an insufficient emotional engagement early in life with the subjective states of others, children with autism struggle later in life to identify what those states might even be. And many think that this impairment is what explains the peculiar sort of struggle that people with autism appear to have with moral judgment.

For instance, Cathy Grant and colleagues found that, while children with autism were as likely as children without autism to judge that damaging *people* makes a person more culpable than does damaging *property*, the justifications provided by children with autism for their judgment were typically of a rather poor quality (e.g., they rarely made any mention at all of such things as the victim's suffering, etc.).⁴¹ In fact, many think that the apparent ability of people with autism to correctly distinguish between moral and conventional wrongs (they often do quite well on Blair's moral-conventional task) has

³⁹ Hobson 2005: 192.

⁴⁰ Hobson 2005: 189, 190.

⁴¹ Grant et al. 2005.

more to do with their having memorized certain rules and other things that they were explicitly taught than it does with an ability to actually *see* a difference between the two.⁴² As Blair explains, “[T]he child with autism may not process *as moral* those transgressions (e.g., stealing) that require an individual to represent the mental state of another *in order to realize that the other is a victim.*”⁴³ An autistic child, then, applies a concept like *morally wrong* in much the same way that our completely colorblind character from earlier applies a concept like *red*—mostly correctly, but nonetheless in a way unlike that of the normal (moral) judge.

Just as we could see the causal significance of Peyton Manning to the Indianapolis Colts’ success during their unfortunate 2011 season without him, we can see the causal significance of emotion to moral judgment by observing the judgment of those with serious emotional deficits. It is not the case that those with emotional deficits get things wrong, as it were—on the contrary, they seem to get things right more often than not. Rather, it is apparently the case that a systematic deficit in emotion forces a person to take some abnormal route to correct moral beliefs, much as the colorblind person must take an abnormal route to correct beliefs about the colors of things. This supports the psychological sentimentalist’s claim that the process by which our moral beliefs are *ordinarily* formed is one in which the emotions figure centrally.

Things now seem doubly bad for the psychological rationalist. For not only do the moral dumbfounding studies suggest that moral intuitions are really the products of a psychological process that operates (mostly) at a level *below consciousness*; but also,

⁴² See Takeda et al. 2007.

⁴³ Blair 1996: 578; emphasis mine. See also Takeda et al. 2007.

there is growing evidence for thinking that this unconscious process is one in which *emotion* figures centrally. I have only discussed some of the relevant empirical research so far, but I nonetheless think that I've said enough to make a strong presumptive case for psychological sentimentalism. In the next section, I'll examine what I take to be sentimentalism's most serious rival.

5. Dual-process models of moral judgment

There are some who think that the traditional rationalist vs. sentimentalist debate in moral psychology is a misguided one, since it fails to acknowledge a kind of “middle ground” approach according to which both reason *and* emotion figure centrally in moral judgment. For instance, in recent work, Fiery Cushman, Liane Young, and Joshua Greene have defended a *dual-process model* of moral judgment.

[W]e pursue a dual-process approach in which moral judgment is the product of both intuitive [i.e., fast, effortless, unconscious] and rational [i.e., conscious, deliberative] processes, and it is the product of what are conventionally thought of as “affective” and “cognitive” mechanisms. [... A] dual-process model of moral judgment can explain features of the data that unitary models [like psychological rationalism and sentimentalism] cannot.⁴⁴

In my view, this dual-process approach is the most significant rival to sentimentalism in moral psychology, so I want to use this section to examine a few of the arguments in favor of dual-process models, and whether or not they really do a better job than sentimentalism of explaining certain phenomena associated with moral judgment. Cushman, Young, and Greene describe four reasons for thinking that reason—or some

⁴⁴ Cushman et al. 2010: 48-49. Readers should note that while it is unclear exactly *how* the authors are using the term ‘moral judgment’ here, it is clear that they are *not* using it in the same way that I have been using it. As I use the term, moral judgment is not the *product* of any psychological process; rather, moral judgment *is* the psychological process that produces moral intuitions. If I were to rephrase this quotation, then, I would say, “Moral judgment *involves* both intuitive and rational processes, and also what are conventionally thought of as ‘affective’ and ‘cognitive’ mechanisms.”

non-emotional, cognitive process—*also* figures centrally in moral judgment, as the dual-process model proposes: (i) the fact that people often *are* capable of offering principled justifications for their moral intuitions; (ii) the fact that people sometimes alter, or abandon, their initial moral intuitions when prompted to engage in conscious reasoning; (iii) the astonishing regularity in the development of moral justifications, as revealed by Kohlberg’s research; and (iv) evidence of heightened activity in areas of the brain associated with higher cognition during some cases of moral judgment.

The first two reasons, I think, are quite weak. First, while cases of moral dumbfounding—i.e., cases in which people are *not* able to offer principled justifications for moral intuitions they are nonetheless unwilling to give up—are admittedly somewhat uncommon, these cases are not supposed to be interesting because of how common or uncommon they are. Rather, they are interesting because of what they might reveal *about* the more common cases in which people can apparently easily justify their moral intuitions—namely, that in these more common cases, it is not that reasoning played any role in the *production* of the intuitions; it is just that, in these cases, reasoning is successful in its *post hoc* rationalization of a moral intuition that has already been formed by some unconscious process. So I hardly see how the fact that people are often able to provide principled justifications for their moral intuitions is any evidence for thinking that those intuitions are the *products* of a psychological process in which reasoning (conscious or unconscious) figures centrally.

Likewise, to say that conscious reasoning can sometimes lead people to alter, or even abandon, their initial intuitions is to say nothing at all about the psychological

process that produced the beliefs in the first place. This is still only to allow reasoning a *post hoc* role with respect to our moral intuitions. It is not to allow reasoning any role at all in *moral judgment*, as I am using that term.⁴⁵ So these first two reasons given by Cushman, Young, and Greene are not problems for the psychological sentimentalist.

What are we to say, though, about the apparent regularities uncovered by Kohlberg in the sorts of justifications that people offer for their moral intuitions? Even if one does not think that his methods were fit for his stated purposes, surely there is something to be said about *how well* his data fits his overall theory. A thorough response to Kohlberg's program would take much more space than I'm prepared to give it here, so my remarks in response to this point will have to be somewhat superficial. Ultimately, I think Kohlberg *has* given us a theory of *something*, but it is a theory of moral *justification* and not a theory of moral *judgment*. And as such, it is supported quite impressively by his research. Again, as I am using the term, moral judgment is the psychological process that underlies the production of a certain type of belief in ordinary cases. Moral justification, on the other hand, is every bit as much a social phenomenon as it is a psychological phenomenon, if not more so. It is the process by which we *make our beliefs understood*, both to others and also to ourselves. At its heart is the giving of reasons, for the purposes of both defense and persuasion. To engage in moral

⁴⁵ To be sure, there is still plenty of room for one to believe that reasoning (conscious or not) has a very significant role to play in the shaping of our *overall moral views*. It is compatible with everything that I say in this dissertation, for instance, that moral judgment (as I define the term) is only *one of many* moral doxastic practices in which humans regularly engage. And reasoning may have a very significant role to play in those other processes by which moral beliefs are formed, shaped, and solidified.

For reasons that I describe in chapter I, however, I do think that moral judgment has a certain *prominence* among other moral belief-forming processes. In most cases, I think, even when we *do* reason about moral matters, we do so *from* our moral intuitions—they are, in a sense, the starting-point of moral reasoning. And it is probably not a coincidence that, much more often than not, when we engage in moral reasoning, we tend to find *more reasons* for thinking that our initial moral intuitions were on point to begin with. But this is only speculation, and I'll not pursue the matter any further than this.

justification, then, is to engage in something that not only rational, but *essentially* rational. It is a kind of skill, and like many skills, it develops over time, as we are exposed to new experiences, new sources of information, new methods, new standards, etc. This is what I think Kohlberg spent the better part of his career studying, not moral judgment.

Now, there is still the issue of explaining just why it is that people's moral intuitions—*however* they are formed—seem so nicely to fit with the sorts of theoretical considerations that are supposed to be characteristic of moral judges at each stage of Kohlberg's model. Why, for instance, do so many people at stage 5 have moral intuitions that correspond so nicely with Kohlberg's description of stage 5 moral reasoning, if these intuitions are really the products of a process in which emotion figures most centrally? There are two ways for a sentimentalist like myself to respond here, one cheaper than the other. The cheap response is just to say that the correspondence was self-fulfilling. As a matter of fact, there is a huge amount of overlap in the actions proscribed or prescribed by different moral theories. If theories A, B and C all forbid actions of types *x*, *y*, and *z*, and I conduct a series of moral judgment surveys in which subjects judge *x*-, *y*-, and *z*-type actions to be wrong, then adherents of A, B, and C are all likely to interpret this as evidence that people are, by and large, guided in their moral judgment by the principles of theories A, B, or C. Kohlberg was an avowed Kantian, so it is not surprising that he interpreted people's intuitions as conforming to Kantian principles of morality. A utilitarian might see the data quite differently.

So much for the cheap response. A more interesting response, I think, might take us back to my theory of emotion from the last chapter. As you will recall, on my account, whenever someone has an emotion, she construes something in a way that clashes or accords with her *concerns* (which then gives rise to the sorts of feelings that we typically associate with emotions). Now, suppose we grant that, as a matter of fact, many relatively mature moral judges are concerned about things like respect for persons, equality and fairness. (We do not have to suppose that they are aware of this fact about themselves.) If this is the case, then presumably, these people will have negative emotional responses whenever they construe an action in a way that clashes with one or more of these concerns. The resulting moral intuitions, then, will likely *fit with* Kantian principles of morality on many occasions, even if they were not the *products of* a reasoning process in which Kantian principles are applied to particular cases, as Kohlberg believed. So in a sense, this person would be emotionally sensitive to violations of Kantian moral principles, but she would not be sensitive to them *as* violations of Kantian moral principles. Now, this is admittedly still a sketchy response; but I nonetheless think that a sentimentalist can satisfactorily explain Kohlberg's data along roughly these lines. And as long as such an explanation is available to the sentimentalist, I do not think that Kohlberg's data should count as evidence for a dual-process model of moral judgment—i.e., *as opposed to* psychological sentimentalism.

Finally, we have to consider the neuroimaging studies that implicate activity in regions of the brain associated with non-emotional, higher cognitive function as subjects engage in moral judgment. In 2001, Greene and colleagues published the results of a

study in which subjects were asked to respond to two of the famous trolley scenarios.⁴⁶ In the first scenario—TROLLEY—a runaway trolley is headed for five people who are, for whatever reason, unable to remove themselves from the trolley’s path. If the trolley continues unimpeded on its current course, it will surely kill all five of them. You stand next to a switch that, if pulled, will redirect the runaway trolley onto an alternative track where it will strike and kill only *one* person. When asked whether or not it is morally permissible to pull the switch and divert the trolley onto the alternative track, sparing five lives at the cost of one, the vast majority of people respond that it is indeed permissible to pull the switch. In the second scenario—FOOTBRIDGE—you now stand on a bridge that spans the trolley track somewhere between the oncoming runaway trolley and the five people on the track. With you on the bridge is a very fat man. You know that if you were to push him down onto the track, his body is large enough that it would stop the trolley before it strikes and kills the five people. But the fat man will surely die as a result of being hit by the trolley. When asked whether or not it is morally permissible to push the fat man off the bridge, sparing five lives at the cost of one, the vast majority of people respond that it is *not* morally permissible to push him.

The puzzle should be clear: Why do so many people judge it permissible to spare five lives at the cost of one when doing so only involves pulling switch, but *not* when doing so involves pushing someone off a bridge? The puzzle becomes perhaps even more troublesome when we consider the sorts of justifications that people provide for their responses. Whenever I present TROLLEY to my students, and then ask them why they think it’s permissible to pull the switch, the predominant justification is markedly

⁴⁶ Greene et al. 2001.

consequentialist: Because five deaths is a worse outcome than just one death. But if a consideration of this basic form is what lead to their response that it is morally permissible to pull the switch, why did it not also lead them to think it permissible to push the fat man—after all, five deaths is *still* worse than just one death, right? In fact, the justifications that people often offer for their response to FOOTBRIDGE are typically more *deontological* in nature, often referring to such things as the man's rights, or a duty to refrain from using people in this way.

Philosophers have attempted all sorts of principled defenses of people's intuitions in response to TROLLEY and FOOTBRIDGE. (Notice, apparently, the assumption is that the intuitions themselves are not the problem; rather, the problem is that of explaining why it is *in fact* permissible to pull the switch, but not to push the fat man. Here is an instance of philosophers granting a certain measure of authority to people's ordinary moral intuitions, as I mentioned in chapter I.) One way to do so, for instance, is to point out the Kantian idea that it is impermissible to use a person as a mere means to some end. The one person on the alternative track is not being used as a means to saving the five, but the fat man surely is. And indeed, anecdotally, I have occasionally had students offer something like this defense of their intuitions in response to TROLLEY and FOOTBRIDGE. So maybe people's intuitions here are the results of a process that is sensitive in some way to exactly the sorts of considerations that Kant would say are relevant to our thinking about the rightness or wrongness of pulling the switch and pushing the fat man.

Greene is not concerned to *defend* people's responses to the trolley scenarios, but he does want to *explain* why they respond in the confusing ways that they do. And here

the explanation that he offers may be surprising to some: Contrary to what is often assumed, deontological moral intuitions (e.g., the majority response to FOOTBRIDGE) are actually the causal result of people's emotional responses, while consequentialist moral intuitions result from a more cognitive, and emotionally neutral, process. The implication is that, if only emotion were not to interfere, subjects would judge it permissible to push the fat man. But the thought of heaving someone off of a bridge to his certain death triggers a kind of "alarm bell" emotional response, which then causes people to judge it wrong to push him. Clearly, if the evidence confirms Greene's hypothesis, this would lend support to his dual-process model of moral judgment against the sentimentalist's model, since it would imply that moral judgment is a process in which emotion only *sometimes* figures centrally—i.e., whenever it competes with the default, cognitive process. So what evidence is there?

Greene's hypothesis generates the following two predictions. First, we should expect to see areas of the brain associated with emotion especially active as subjects consider FOOTBRIDGE, and areas of the brain associated with higher cognition especially active as subjects consider TROLLEY. And second, when it comes to people's responses to FOOTBRIDGE in particular, we should expect to see a significant difference in reaction time between those who judge it impermissible to push the fat man and those who judge it permissible. After all, if the majority response to FOOTBRIDGE is the causal result of an automatic emotional reaction, then those in the minority who judge it permissible to push the fat man will presumably have to override this emotional reaction, and this overriding process will take time.

Both of these predictions have been confirmed. The fMRI scans show increased activity in the posterior cingulate cortex, the medial prefrontal cortex, and the amygdala—three areas closely associated with emotion—as subjects consider FOOTBRIDGE, and increased activity in the dorsolateral prefrontal cortex and inferior parietal lobe—areas closely associated with more “cognitive” functions—as subjects consider TROLLEY.⁴⁷ Furthermore, on average, subjects who approve of pushing the fat man take much longer to do so than subjects who disapprove of pushing him.⁴⁸ So the evidence does seem to support Greene’s hypothesis that moral intuitions are *sometimes* the causal products of a process in which emotion figures centrally, and, other times, the products of a more emotionally neutral, “cognitive” process. Unless the sentimentalist wants to claim that people’s intuitions about TROLLEY are somehow out of the ordinary—and so, not spoken for by his account of moral judgment—he will have to offer some sort of counter-interpretation of Greene’s studies.

Here is how I think such an interpretation should go. When two or more emotional responses pull moral judgment in different directions, our brains recruit more neutral, cognitive processes to help adjudicate. Otherwise, the “default” setting for moral judgment is just to let emotion take the lead (contrary to what Greene suggests). If this is right, then here is what we should expect to see. In cases in which emotion is, for whatever reason, unable to produce a stable moral intuition on its own (e.g., because we have comparably strong emotional responses that lend themselves to competing moral intuitions), there will be increased activity in areas of the brain associated with higher

⁴⁷ Greene et al. 2001: 2106-2107.

⁴⁸ Greene et al. 2001: 2107.

cognition, and reaction times will generally be relatively high. And in cases in which emotion has no difficulty producing a stable moral intuition on its own, there will be less activity in those cognitive areas of the brain, and reaction times will be significantly lower.

This is what we observe in the cases of TROLLEY and FOOTBRIDGE. In response to the TROLLEY scenario, emotional responses to the thought of pulling the switch, though perhaps relatively weak, are close enough in strength to subjects' emotional responses to the thought of letting the trolley continue on its course; and so, more cognitive processes are recruited to locate some consideration that might "tip the scales," as it were—e.g., the thought that five deaths would be worse than just one death. In response to the FOOTBRIDGE dilemma, however, subjects' emotional responses to the thought of pushing the fat man are strong enough to tip the scales on their own; so, as far as our brains are concerned, no help is needed in forming the intuition that it would be wrong to push him.

Now consider two other cases given to subjects by Greene and his colleagues.

Call the first one CRYING BABY:

It is wartime, and you and some of your fellow villagers are hiding from enemy soldiers in a basement. Your baby starts to cry, and you cover your baby's mouth to block the sound. If you remove your hand, your baby will cry loudly, the soldiers will hear, and they will find you and the others and kill everyone they find, including you and your baby. If you do not remove your hand, your baby will smother to death. Is it okay to smother your baby to death in order to save yourself and the other villagers?⁴⁹

The second case—INFANTICIDE—involves a teenage mother who must decide whether or not to kill her unwanted newborn. As Greene reports, in the case of CRYING BABY,

⁴⁹ Greene 2008: 44.

“Different people give different responses, and nearly everyone takes a relatively long time.”⁵⁰ But in the case of INFANTICIDE, subjects are quick and unanimous in their response that it is not permissible for the young mother to kill the unwanted newborn.

Here again, while our negative emotional response to the thought of the teenage mother killing the newborn is strong enough on its own to cause us to have the intuition that doing so is impermissible, the CRYING BABY case is much more likely to give rise to an emotional conflict. And since emotion is unable in this case to produce a stable moral intuition on its own, our brains recruit some more cognitive process in order to locate considerations that might “tip the scales.” The result *might* be something consequentialist (e.g., “I suppose it’s better for just one person to die than for that same person and many others to die as well”), but it may also be something more deontological (e.g., “I don’t care if it means everyone has to die, it would be wrong of me to use the baby as a means to their survival”), or something unrecognizable as an instance of either substantive ethical framework (e.g., “I just could never do something like that”). As Greene said, responses were varied.⁵¹ My point is only that, while moral judgment is, by default, driven by emotion, there will nonetheless be cases in which emotional responses alone are indecisive. And in such cases, it should not be surprising that our brains recruit “outside help,” as it were. In fact, one of the regions of the brain that is especially active as subjects consider CRYING BABY is the anterior cingulate cortex, which, as Cushman, Young, and Greene explain, “reliably responds when two or more incompatible

⁵⁰ Greene 2008: 44.

⁵¹ So this is a case in which the data do not support Greene’s contention that there is some relationship between higher cognitive function and the production of characteristically *consequentialist* moral intuitions.

behavioral responses are simultaneously activated, i.e., under conditions of ‘response conflict’.”⁵²

This is all perfectly compatible with psychological sentimentalism. In fact, sentimentalists of all stripes have generally been quite happy to allow that there may be some roles for non-emotional processes to figure in the production of our moral beliefs. They just insist that these roles are relatively *minor* compared to that played by emotion. As Hume famously remarked, “Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them.”⁵³ Psychological sentimentalists carry the Humean torch, in a sense, echoing the conclusion of Schnall, Haidt, and others: “[M]oral reasoning is an important part of moral life, but for most people, most of the time, most of the action is in the quick, automatic, affective evaluations they make of people and events.”⁵⁴

6. Is psychological sentimentalism an epistemic threat?

The phenomenon of moral dumbfounding suggests that our moral intuitions are ordinarily the products of a psychological process that operates (mostly) at a subconscious level, contrary to what psychological rationalists believe. And the research discussed in §§3-5 strongly suggests that this psychological process is one in which emotion figures centrally. Henceforth, I will assume that some form of psychological sentimentalism is the case. What sort of role does emotion play in moral judgment? Apparently, a very *significant* one. In the next two chapters, my primary concern will be

⁵² Cushman et al. 2010: 51.

⁵³ Hume 2000 [1739/1740]: 266.

⁵⁴ Schnall et al. 2008: 1097.

the exact *nature* of the role that emotion plays in moral judgment: Is emotion a distraction to moral judgment? Does it cloud moral judgment? Or might emotion's contribution to moral judgment be something more positive, epistemically speaking?

As I mentioned in chapter I, there are some who argue that the centrality of emotion in moral judgment is somehow a threat to the epistemic status of moral intuitions (or *would be* a threat, if they also believed psychological sentimentalism to be the case). Generally speaking, underlying these arguments is a long-standing, and still very popular, assumption that *if* emotion figures centrally in moral judgment—as we've now established it does—then moral judgment probably cannot be relied upon to produce justified beliefs about objective moral values. Joshua Greene makes this assumption when he moves *from* the observation that emotion figures centrally in moral judgment *to* the conclusion that our moral intuitions must then be projections of internal attitudes rather than beliefs about objective values.⁵⁵ Sharon Street makes the assumption on the basis of observations about the purpose(s) for which emotions evolved. Shaun Nichols also appears to make the assumption when, after defending his own version of psychological sentimentalism, he concludes, “so moral judgment, as we know it, is not objective.”⁵⁶ And Walter Sinnott-Armstrong must have something like this assumption in mind when he argues that, whenever our moral intuitions are the products of a process in

⁵⁵ Greene 2003. In a recent interview, Greene was asked how he believes the empirical research (i.e., the research that suggests that our moral intuitions are to a great extent based upon emotion) bears upon the justificatory status of people's ordinary moral intuitions. He replied that he thinks it presents a kind of “debunking explanation” of these intuitions, i.e., an explanation of “why he have a belief in a way that makes it unlikely that the belief is true” (Sommers 2009: 137).

⁵⁶ Nichols 2004: 185.

which emotion figures centrally, some further inferential confirmation will be needed if they are to count as justified.⁵⁷

I'll address all four of these arguments in much greater detail in chapter V. First, though, I want to bring the discussions of chapters II and III together, and lay out my own account of the role of emotion in moral judgment and its epistemic implications. That will be the focus of chapter IV.

⁵⁷ Sinnott-Armstrong 2006.

IV MORAL PERCEPTION

[F]ear and confidence and appetite and anger and pity and in general pleasure and pain may be felt both too much and too little, and in both cases not well; but to feel them at the right times, with reference to the right objects, toward the right people, with the right motive, and in the right way ... is characteristic of virtue.

Aristotle, *Nicomachean Ethics*

When I say, “The sky is blue,” and then say, “The sky is beautiful,” a property is attributed to the sky in either case. In the second case a feeling participates in the apprehension of the property, as, in the first case, an idea does.

Alexius Meinong, *On Emotional Presentation*

At the end of chapter I, I made a point to clarify my use of two key terms: *moral intuition* and *moral judgment*. Moral intuitions are a subspecies of moral beliefs, characterized (among psychologists) by the quickness and effortlessness with which they are formed, and (among philosophers) by their non-inferential grounds. And moral judgment is the psychological process that is ordinarily responsible for the production of moral intuitions—in which, it now seems, emotion figures centrally.

Now I want to introduce and define another term, one that will already be familiar to readers, but will have a specialized meaning here. The term that I have in mind is *trustworthiness*, and I’ll apply it to moral intuitions. As I explained in chapter I, we go on as if our moral intuitions are beliefs *about objective moral values*, and as if they are at

least *sometimes justified*—even if, as is typically the case, we never actually consider what grounds *there are*, or what grounds *we have*, for doing so. I’ll say that ordinary moral intuitions are trustworthy, then, if it turns out that they actually do have the characteristics that we assume they have—i.e., if they actually *are* justified beliefs about objective moral values.¹

The central question of this dissertation can now be put in these terms: Can moral intuitions be trustworthy if they are grounded (causally) in emotion? Or, since I’ve already argued that moral intuitions *are* grounded in emotion, we can now simply ask: Are moral intuitions trustworthy? My answer will be that, yes, they *are* (or they *can be*, under conditions I specify in this chapter). But as I’ve indicated elsewhere, this is not a popular answer—especially among philosophers and psychologists. In general, the assumption seems to be that *if* emotion figures centrally in moral judgment, then its products are either (a) *not* beliefs about objective moral values (but rather, for instance, mere projections of emotions), or (b) if they *are* beliefs about objective moral values, then their justification is seriously threatened by their emotional basis. In other words, it is commonly assumed that, other things being equal, moral intuitions *cannot* be trustworthy if they are grounded in emotion.

Something has to give. If psychological sentimentalism is the case—and again, I’m assuming at this point that it is—then we ought either somehow abandon the trust that we apparently place in our own moral intuitions (since apparently they are not *worthy* of our trust), or else reject the popular assumption that our moral intuitions cannot be justified beliefs about objective moral values if they are grounded in emotion. Given

¹ I’ll say more below (in §4) about the sort of epistemic justification that I have in mind.

the sort of role that moral intuitions play in our lives and in the construction and assessment of moral theories, the first of these options would be difficult to accept, to say the very least. And yet, it is precisely what some philosophers and psychologists recommend. In this chapter, I attempt the second response—that is, I’ll offer an account of how it is that ordinary moral intuitions might be worthy of the trust that we apparently give to them.

For now, consider a rough analogy: If my beliefs about the shapes of objects are grounded in my visual experiences of those objects, then, other things being equal, the beliefs are trustworthy (in an analogous sense of trustworthiness). This is relatively uncontroversial because we already have a story to tell about how our sensory receptors are appropriately sensitive to such objective features of the world as the shapes of things. So in order to make such a claim about emotion and morality—i.e., in order to make the claim that moral intuitions grounded in one’s emotional responses can be similarly trustworthy—one will have to tell a story about how it is that our emotions can be (and sometimes *are*) sensitive to objective moral values. Telling such a story shall be the main task of this chapter.

Now, some readers may wonder, *hasn’t this story been told before?* Indeed, John McDowell defends a sensibility theory according to which people might be sensitive, in a manner reminiscent of perception, to objective moral values.² But McDowell’s approach involves—and indeed, on some readings, depends upon—the adoption of a secondary quality view of moral values, according to which a thing’s having a moral value is a matter of its *being such as to elicit a certain response in perceivers*. While McDowell

² See McDowell 1998a and 1998b.

may be right to construe moral values in this way, though, it is obvious that skeptics like Nichols and Sinnott-Armstrong do not—for if they did, it would be difficult to make sense of their skepticism. If *being such as to elicit certain emotions*, say, were all there is to objective moral values, then why think that having one's moral intuitions grounded in emotion is any threat to their trustworthiness? Taking McDowell's secondary quality approach to moral values may be *one* way to protect moral intuitions from the sort of skepticism encouraged by Nichols and Sinnott-Armstrong, but I suspect this would feel like a hollow victory to some, especially those inclined to construe objective moral values more robustly than this. And apparently, taking such an approach would amount to simply refusing to operate on the terms assumed by these skeptics.

What I propose to do, then, is this: *Assume*—just for the sake of argument—a primary quality view of objective moral values, and *then* raise the epistemological question: What becomes of the trustworthiness of beliefs about moral values so construed, when it is granted that these beliefs are grounded in emotion?³ My answer will depend heavily upon the account of emotion that I defended in chapter II, and upon an account of the relation between emotion and virtue that I will describe in §1 of this chapter. In §2, I return to the analogy to perception, arguing that the effect of moral virtue upon the emotions is such as to make the virtuous person emotionally sensitive to objective moral values in a manner much like perception—if you prefer, we might say that the virtuous person's emotions amount to instances of *moral perception*. And just as the perceptual experiences in which our perceptual beliefs are grounded (can) contribute

³ To be clear, I take myself to be free to assume a primary quality view of objective moral values precisely because nothing I will say depends upon how objective moral values are construed. As far as I can tell, everything I say in this chapter can be affirmed by someone who holds a secondary quality view like McDowell's. The point is just that we *needn't* hold such a view.

positively to the justification of those beliefs, our emotions can make a similarly positive contribution to the justification of moral intuitions. Other things being equal, I conclude, if a person is virtuous, then her moral intuitions are trustworthy *because* they are grounded in emotion, contrary to what is commonly assumed.

In §3 and §4, I elaborate upon my view. In §3, I do this by distinguishing it from a somewhat similar view defended very recently by Robert Roberts. Roberts also argues that emotion can be a kind of epistemic aid to moral intuitions, but I think there are important differences between the ways in which he and I account for this. And since my view can be understood as a form of reliabilism with respect to moral intuitions, I devote §4 to a brief discussion of the view in the broader context of process reliabilist accounts of epistemic justification. Finally, in §5, I address a few lingering worries.

1. Why do good feelings happen to good people?

Crucial to my overall argument that moral intuitions can sometimes be trustworthy when they are grounded in emotion is the establishment of a kind of connection between emotion and *moral virtue*. It would certainly be nothing new to suggest that emotion and moral virtue are somehow related. As Linda Zagzebski notes, “[A]lmost every writer on the moral virtues has connected them with feelings.”⁴ Perhaps most famous among them is Aristotle and his remarks about the connection between emotion and virtue in the *Nicomachean Ethics*. Now, I should note that Aristotle seems to have thought differently about the nature of emotion than I do, and would consequently not attribute to emotion anything like the role in moral judgment that I will in this dissertation. But for now, that is beside the point. All I want to focus on, for the moment, is the basic sort of

⁴ Zagzebski 1996: 128.

relationship that Aristotle thought existed between emotions and moral character. In the *Nicomachean Ethics*, Aristotle describes this relationship as one in which character plays some role in determining the emotions, so that our emotional responses are at least often signs or symptoms of the state of our moral character. He writes,

[F]ear and confidence and appetite and anger and pity and in general pleasure and pain may be felt both too much and too little, and in both cases not well; but to feel them at the right times, with reference to the right objects, toward the right people, with the right motive, and in the right way [...] is characteristic of virtue.⁵

So, to be morally virtuous is (partly) to be disposed to having certain emotional responses; and at least generally, the emotional responses of the morally virtuous person will be better than those of the less virtuous. Throughout what follows, I'll assume that this broadly Aristotelian picture of the relation between emotion and moral character is correct.

Why is it that morally virtuous people feel emotions well? As I see it, there are basically two approaches that one might take to giving an answer to this question. The first is certainly the easier of the two, but it is also explanatorily less satisfying: Simply *define* virtue in terms of having the right emotions. The virtuous person feels emotions well, then, because *that's just (part of) what it is to be virtuous*. The problem with taking this approach, however, is essentially the same as the problem of trying to forge a conceptual link between virtue and right action: Responding emotionally to the right things, in the right ways, etc., is neither necessary nor sufficient for being virtuous (just as acting rightly is neither necessary nor sufficient for being virtuous). It is not necessary because even the most virtuous person can occasionally fail to respond emotionally in the

⁵ Aristotle 1106b17–23

ways that she should, perhaps because she is distracted, or ill-informed, or whatever. And it is not sufficient because someone who responds emotionally to things in all the right ways, but always ignores or resists those emotions, would hardly count as virtuous. So an answer to the question posed in the heading of this section—why do good feelings happen to good people?—just cannot be as simple as: Because that’s what it is to be good. But then, what would a satisfactory answer look like?

The other approach to an answer is decidedly more difficult, but explanatorily more satisfying: Give an account of what it is about virtue, specifically, that disposes a virtuous person to feeling emotions well. In other words, we take the observation that moral virtue disposes a person to appropriate emotional responses as a datum to be *explained by* a theory of virtue, and not as a *part of* our theory of virtue. In taking this sort of approach, one takes on an explanatory burden that the first approach bypasses; but it is a burden that I nonetheless think that I can shoulder, with the help of my account of the nature of emotion.

On the view that I defend in chapter II, some concern or other is necessary if one is to experience an emotion. If I somehow ceased being concerned about my own safety, for instance, then I’d no longer fear the approaching animal, even if I continued to see it as something likely to harm me. To expound a bit upon my earlier description, I use ‘concern’ in this context as a name for a set of ways of being oriented toward something. I might *love* or *hate* *x*, *like* or *dislike* *x*, *admire* or *despise* *x*, *think highly of* or *think poorly of* *x*, and so forth. If I am oriented toward *x* in any of these ways, it would be correct (given the way that I am using the term) to say that I am *concerned* about *x*, that I have *x*

as an object of concern. And all of these ways of being oriented toward x dispose a person to feel certain emotions toward or about x . If I admire a person, for instance, and then come to see him as the perpetrator of some gross moral evil, then I will likely feel disappointment toward him. (Think, for instance, of the responses of fans of the late Penn State head football coach Joe Paterno to the recent news of his alleged involvement in covering up a former coach's crimes.) In this way, having concerns disposes a person to certain emotions.

Furthermore, we can be concerned about things in different ways and to varying degrees. An investor in some company and an opponent of that company are both concerned about the company's success, but obviously in very different ways. And it is for this reason that the two are likely to respond emotionally in very different ways to news of the company's success or demise. Likewise, two fans of the same baseball team might nonetheless have quite different responses when the team wins a pennant, depending upon the degree of their concern for the team and its success.

So having concerns disposes a person to certain emotions, and concerns can be had toward things in different ways and to varying degrees. I suspect that this will strike the reader as mostly common sense, but it is important to keep in mind, for: (a) as I explained in chapter II, the fact that concerns have these features underlies the susceptibility of our emotional responses to moral evaluation (and perhaps other sorts of evaluation as well); and (b) as I will now go on to show, these features of concerns allow me to explain why it is that emotion bears the sort of relation to virtue and moral character that it does on the Aristotelian picture described above.

Moral virtue, I submit, is a matter of being concerned for the good.⁶ I use the term ‘the good’ here as a kind of blanket term for all that is, in fact, good. So I am not assuming any sort of metaphysical unity to goodness. Rather, I think there are many goods, some of which are in some sense objective. (Again, we can assume, for the sake of argument, that these goods are objective in a primary quality sense; but nothing that I will say here depends upon this.) Other things being equal, a person is morally virtuous to the extent that she is concerned about these things, i.e., to the extent that she has aligned her concerns with the good. She loves, likes, and admires what is good; she thinks highly of the good; she encourages it, protects it, stands for it, and so forth.⁷

⁶ Since I don’t think it’s important, given what I’m up to in this dissertation, I’ll leave open the issue of whether there might be more to virtue than this. So, being concerned for the good may not be *sufficient* for moral virtue, but I do think it is *necessary*.

Also, I should note that, when I say “being concerned for the good,” I mean this in a *de re* sense. Nomy Arpaly makes a similar point in her book *Unprincipled Virtue*, in which she is concerned to develop a new theory of moral praiseworthiness (and blameworthiness). Arpaly understands moral praiseworthiness partly in terms of responsiveness to moral reasons, and as she explains,

I take a person to be *responsive to moral reasons* to the extent that she wants noninstrumentally to take courses of action that have those features that are (whether or not she describes them in this way) right-making and not to take courses of action that have those features that are (whether or not she describes them in this way) wrong-making features.

(Arpaly 2003: 79; italics in original)

In other words, what is most important when it comes to our evaluations of moral worth is not that agents are responsive to *what they see as* moral reasons (or even what they *could* see as moral reasons), but rather that agents are responsive to *what really are* moral reasons. This is why, she thinks, Huck Finn may be praised for his failure to turn in his slave friend Jim. As she reads the story, Huck counts as a praiseworthy agent because, despite his professed beliefs to the contrary, he is nonetheless responsive to moral reasons in a *de re* sense. In this particular case, he is responsive to the fact that Jim is a person, even if he is unaware of this. (It is tempting to cash this out in terms of *construals*: unconsciously, and as a result of the time he spends with Jim, Huck has come to see Jim as a person.) The virtuous person is one who is excellent in being for the good in a similarly *de re* sense—that is, not in being for what *she takes to be* the good, but in being for what *is in reality* the good. I’ll return to this issue briefly in §5.

⁷ For a full defense of a theory of virtue much like the one I describe here, see Adams 2006. According to Adams, moral virtue consists in what he calls “persisting excellence in being for the good.” And what he means by ‘being for the good’ is something very similar (if it is not identical) to what I have in mind when I say that a virtuous person is someone who is concerned for the good. As he explains, “There are many ways of being for something. They include: loving it, liking it, respecting it, wanting it, wishing for it, appreciating it, thinking highly of it, speaking in favor of it and otherwise intentionally standing for it symbolically, acting to promote or protect it, and being disposed to do such things” (Adams 2006: 15–16).

Now, as I noted in chapter II, there are ways of getting things wrong, morally, when it comes to one's concerns—we can be concerned about the wrong sorts of things, for instance, or we can be concerned about the right things but to a wrong degree. My two examples there were: (1) a person who desires to see others humiliated, and is consequently delighted or amused when he construes another as humiliated; and (2) a person who, due to a lack of self-respect, fails to be indignant upon being treated unfairly. Both of these people, I can now explain, represent different types of failures of moral virtue. The first person represents a case of failing to be concerned for the good: other things equal, to want to see others humiliated is to want something that is altogether *bad*. And this is why we would not only disapprove *of the delight* that this person takes in another's humiliation, but further, we would disapprove *of the person* who delights in such a thing. The delight reflects poorly upon the person because it manifests a defect in the person's moral character.

The second person represents a different sort of failure. Presumably, this person is concerned for his own dignity to *some* degree—just, apparently, to an insufficient degree. That is, there are surely some offenses against his dignity that he would not tolerate, but being seriously mistreated by his boss in this particular way happens not to be one of them. It *should* be, though, and that is why we not only disapprove of his lack of indignation, but also of what its absence reveals about his character—namely, a deeper failure to be concerned for his own dignity, his own good, to the degree that he ought.⁸

⁸ On this point, I take myself to be in partial agreement with Aristotle when he writes, “Now neither the virtues nor the vices are passions, because we are not called good or bad on the ground of our passions, but are so called on the ground of our virtues and vices, and because we are neither praised nor blamed for our passions (for the man who feels fear or anger is not praised, nor is the man who simply feels anger blamed, but the man who feels it in a certain way), but for our virtues and our vices we are praised or

Here we see that it is not enough to merely be concerned for the good in *whatever* way and to *whatever* degree. Rather, moral virtue requires a kind of *excellence* in one's concern for the good.

Why is it, then, that our emotions bear the sort of relation to moral character that they do on the Aristotelian picture? Because moral character consists in a particular set of persisting concerns, and having concerns disposes people to certain emotions. In ordinary cases, at least, a person will be saddened by another's death, joyed by news of a friend's recovery from illness, and frightened by a dangerous animal in part *because of her concerns*. And to the degree that her concerns are had *excellently* and *for the good*, hers is a virtuous character, one that disposes her to emotional responses that are better, in some sense, than those of someone less virtuous. This is why, as Aristotle put it, feeling emotions "at the right times, with reference to the right objects, toward the right people, with the right motive, and in the right way [...] is characteristic of virtue."

Furthermore, if I am excellently concerned for the good, then I'll also tend to *construe* things in ways relevant to goodness. After all, in most cases, the features of a person's circumstances that are most salient to her will to a large degree be determined by her concerns. We get annoyed when we are awakened in the middle of the night by loud parties outside because we are concerned both for our sleep health and for the consideration that we feel we're owed by others. And it is because of these concerns that the fact that they are waking us in the middle of the night—as opposed to the fact that,

blamed" (Aristotle 1105b29–1106a2). Unlike Aristotle, I do think that we can legitimately be praised and blamed for our emotions (for more on this point, see Roberts 1989: 295ff.). But like Aristotle, I think praise and blame in this context often imply a deeper sort of evaluation, i.e., evaluation of the character underlying the emotions.

say, they are singing “Copacabana” again—will seem especially salient to us. But if an *intruder* wakes us in the middle of the night, it will not be annoyance that we feel, but fear. Since we are generally more concerned for our survival than we are for things like sleep health and the consideration of others, we construe the intruder not *as someone who interrupts our sleep*, but rather *as someone who threatens our survival*. In these sorts of ways, our concerns affect the way we see things. So if moral virtue is ultimately a matter of having the right concerns, then the morally virtuous person will see things in ways relevant to that for which she is concerned—namely, *the good*. Because of her concerns, certain features of her circumstances will stand out to her, and they will be features relevant to the possession and exercise of virtue.⁹

So, to answer the question posed in the heading of this section—why do good feelings happen to good people?—we can now say the following. First of all, concerns in general dispose us to certain emotional responses; and ‘good people’, in this context, are those whose concerns are aligned with the good. And further, virtue can also influence the way we construe things, since our concerns often determine which features of our circumstances are most salient to us.

How is all of this relevant to moral judgment? And in particular, how is it relevant to the trustworthiness of the products of moral judgment—namely, ordinary moral intuitions? I answer these questions in the next section.

⁹ This is the sense in which I think McDowell is right to suggest that the virtuous person is “one who sees situations in a certain distinctive way” (McDowell 1998a: 73).

2. Being good and feeling (the) good

In *Unprincipled Virtue* (2003), Nomy Arpaly argues that moral psychology is undernourished because it has been feeding on what she thinks is an unbalanced diet of examples. When they are not considering cases “involving bizarre, ill-understood individuals, artificial situations of the kind created by the television show *Survivor*, or horrible moral dilemmas that verge on the grotesque,” Arpaly complains, moral psychologists tend to focus their attention upon “[c]ases in which the mind of the agent seems quite transparent to her, cases in which she either acts for reasons that she understands and endorses or is carried to action by some atavistic force, [...] and cases in which, even when she is irrational, she knows she is.”¹⁰ This is unfortunate because a great deal of our ordinary (moral) psychological lives seems not to be getting the sort of attention it deserves. Lost in plain sight, for instance, are cases of ego-dystonic desires and emotions—i.e., desires and emotions that are seriously inconsistent with one’s self-image—and cases of what she calls “inverse akrasia”—i.e., cases of people doing the right thing against their “best” judgment.¹¹ These are phenomena that Arpaly thinks are actually rather common, but you would never guess this given the amount of attention they are typically given by moral psychologists.

¹⁰ Arpaly 2003: 30, 29.

¹¹ As Arpaly defines it, a person’s “best” judgment is “the judgment that one reaches, having taken into account all of the reasons one judges to be relevant, as to what would be best for one to do in a given situation” (2003: 35).

Also, for the record, in saying that these phenomena are “lost in plain sight,” I do not mean to say that *no one* has noticed them. Rather, I only mean to imply that the frequency with which they occur in real life vastly exceeds the frequency with which they are discussed in moral psychological literature. These phenomena are much more familiar to our ordinary lives, for instance, than things like zombies, aliens, and evil demons; and yet, for whatever reason, one is far more likely to find moral psychologists discussing these latter sorts of cases.

Another phenomenon that Arpaly suggests deserves more attention is that of “rationality without deliberation,” which can apply to both action and belief formation. Given that my interests here are epistemological rather than action theoretic, I shall focus on cases of rational belief formation without deliberation, i.e., cases in which beliefs arise out of a process that does not involve any kind of *reflection* upon one’s reasons for belief, but that is nonetheless *responsive* to those reasons in a way that confers some degree of justification upon the resulting belief. Consider, for instance, cases of what Arpaly calls “dawning.” These are cases in which people arrive at some belief “sans deliberation, as a result of a long period of exposure to new evidence”:

Candide, a young man who lives in an idyllic setting, accepts the authoritative Dr. Pangloss’s view that this is the best of all possible worlds, in which everything is for the good, humankind is nice, and so on. [...] Then, after gaining some experience with the wider world, Candide still swears that this is the best of all possible worlds, but there is less conviction in his voice, or perhaps there is an excess of conviction in his voice (as he protests too much) but his face is no longer that of the wholehearted optimist. Then one day, when asked if he believes that this is the best of all possible worlds, he discovers that he does not think so and has not for some time. Here we have a process, starting with firmly believing that this is the best of all possible worlds and ending with disbelieving the same proposition. At no point in the process, let us suppose, did deliberation occur [...] But we do not regard Candide’s change of mind as irrational. In fact, we regard it as rational. Finally, the young man came to his senses.¹²

Two things are worth noting. First, dawning is a relatively common phenomenon. This is often how people describe, for instance, (i) conversions into, or out of, some moral, political, or religious view, (ii) the realization that a friend or co-worker has become an object of romantic love (or perhaps, the realization that one has fallen *out* of love with another), and (iii) the discovery of unfortunate psychological scars from past

¹² Arpaly 2003: 54.

experiences.¹³ And second, the fact that these newfound beliefs (sometimes) arise out of a process that involves exposure to evidence for their truth is important, for it suggests that, other things being equal, the new beliefs may be on a better epistemic footing than their alternatives. In response to the question, “Is Candide justified, after years of ‘experience with the wider world’, in believing that this is *not* the best of all possible worlds?” it is surely right to say that he is at least *more* justified in believing that than he would be in maintaining his former belief. Even if we do not want to say that he is *as justified as he could possibly be*, it does seem right to say that the process out of which his newfound belief arose is one that *contributes positively* to its justification. After all, it was formed on the basis of evidence, even though this fact about the belief was not itself transparent to Candide in the meantime.

An even more common case of rational belief formation without deliberation is that of ordinary perceptual belief. The epistemology of perception is a notoriously thorny subject, but there are two things upon which almost everyone seems to agree. First, ordinary perceptual beliefs are not formed on the basis of any kind of deliberation. It may of course be true that perception depends in some way upon concepts—so that, e.g., I cannot perceive a red cup without first having, and at some point deploying, the concepts *red* and *cup*. But few, if any, think that we can only deploy concepts like this in deliberation. We do form ordinary perceptual beliefs *on the basis of* perception, but not

¹³ For the record, I do not mean to suggest that all such cases of dawning are rational. As Arpaly is careful to note, just as we can deliberate irrationally, beliefs can dawn on us irrationally. Her example is that of someone who describes his conversion from hopeless, despairing loneliness to belief in the existence of Jesus in the following way: “Suddenly, when I was in bed, trying to sleep, it dawned on me that I can’t really be so alone. It is simply impossible for life to be so unbearable; someone must be watching me. Jesus must be watching me” (2003: 56). The difference here, as Arpaly puts it, is that the relationship between the evidence (e.g., her loneliness, and other sources of misery in her life) and the conclusion (that Jesus must be watching her) is not a compelling one.

in the same sense that, for instance, we arrive at conclusions *on the basis of* consideration of premises. I do not think to myself, “Well, I have the perception of a red cup, my perceptual faculties are in general reliable,” and so forth, and then *infer* my way to the belief that there is a red cup before me. Rather, I just see the red cup and then believe that there’s a red cup. In fact, there are some who think that perception is itself sufficient for belief (as they say, “seeing is believing”).

And second, regardless of one’s views on the epistemology of perception, most agree that ordinary perceptual beliefs—i.e., those formed as a result of the normal functioning of our perceptual faculties, in conditions sufficiently optimal for the functioning of the relevant faculty, etc.—despite their non-deliberative basis, are nonetheless on-balance justified. There may of course be *more* that I can do to justify my belief that there’s a red cup before me—e.g., I could confirm that others similarly situated have the same perceptual experience; I could better optimize the conditions in which the perceptual experience is had (brightening the lights in the room, cleaning the lenses of my glasses, and so forth); etc. But even without doing these things, I’m surely justified to some degree in believing that there’s a red cup before me, if the belief is grounded in my perceptual experience of a red cup (and, of course, the experience is had in conditions optimal for the functioning of the relevant faculty). The perceptual experience itself contributes positively to the justification of the relevant belief. And this must have something to do with the nature and function of my perceptual faculties. My visual system, for instance, when conditions are sufficiently optimal for its normal function, interacts with the world in such a way as to give me reason to trust its deliverances. That is, my visual system is sensitive to certain features of reality, and, other things being

equal, this sensitivity goes at least some way toward justifying beliefs made on its basis about those features of reality to which it is sensitive.

Now, in chapter III, I explained that the psychological sentimentalist believes that emotion plays the same sort of role in moral judgment that perception plays in perceptual judgment, i.e., that our moral intuitions are formed on the basis of emotional responses in basically the same way that ordinary perceptual beliefs are formed on the basis of perceptual experiences. I believe that Jerry Sandusky behaved wrongly *because* his behavior horrifies me emotionally, just like I believe that there is a red cup before me *because* I have the perceptual experience of a red cup. I now want to see if I can push this analogy even further, by asking: Just as *perception* can contribute positively to the justification of beliefs formed on *its* basis, might *emotion* contribute positively to the justification of beliefs—specifically, moral intuitions—formed on *its* basis? I think it can, as long as emotion is understood along the lines I describe it in chapter II, and as long as the relation that emotion bears to moral character is roughly as I characterize it in §1 of this chapter.

Consider what I shall call the *maximally virtuous person*, or MVP for short. Along the lines sketched in §1, the MVP is someone who is as excellent as possible in his concern for the good. He has all the concerns that he ought to have, in whatever way, and to whatever degree, he ought to have them. And he is as appropriately sensitive as a person can be, in the way he sees things, to those features of his circumstances that are most relevant to the exercise of moral virtue. Now, suppose that some person S performs some action ϕ , and that the MVP responds to S's ϕ -ing with anger. The details are

unimportant, and in fact, my point here may be better served if we know nothing of the nature of S and ϕ *except* that S's ϕ -ing angers the MVP.

I want to make two observations about this case. First, the fact that the MVP is angered by S's ϕ -ing is itself a good reason for thinking that S's ϕ -ing was (objectively) *morally wrong*. After all, S's ϕ -ing has not angered *just anyone*; it has angered someone who is as excellent as possible in his concern for the good. Here is an analogy. My wife is a skilled musician and classically trained singer who has spent many years studying, composing, performing, and even teaching, all varieties of music. She has what many would call a "musical ear," being incredibly sensitive to such things as pitch, tone, rhythm, melody, and the like. Now suppose she is listening to a vocal performance, and when the singer hits a particular note, the hairs on the back of my wife's neck stand up and she cringes. I take it that my wife's response is as good a reason as any for her to think that the singer was out of tune. After all, the note in question did not make *just anyone* cringe; it made cringe someone who is especially sensitive to breaches of melody, harmony, and the like. Likewise, given the way in which I understand emotion and virtue, the MVP is someone who is especially *emotionally* sensitive to breaches of goodness, and so his response to S's ϕ -ing is as good a reason as any for thinking that S's ϕ -ing was wrong.

In fact, if we can say that my wife not only hears the note, but even that she *hears its flatness*, then I hardly see what should keep us from saying that the MVP *feels the wrongness* of S's ϕ -ing (as long as we assume, as I do here, that there are such things as flatness and wrongness). After all, given the way that I understand the nature of emotion,

and given that the MVP's concerns here are for the good, we can say that he is feeling the contrariety of S's ϕ -ing to the good. And this contrariety is not of the MVP's making. It is a fact about S's ϕ -ing, independent of the MVP's beliefs, concerns, emotions, etc. It is *through* his emotions, though, that he is attuned to these sorts of facts. In this way, it seems emotions at least *can be* sensitive to moral values in a manner much like perception, provided that the subject of the emotions is sufficiently virtuous.

Can we say that a person's emotions are sensitive to objective moral values *precisely to the degree to which* that person is virtuous? That would be nice and simple, but I also think that it would be false. Compare two people with similar lives: each has been through difficult times, the results of which have left each with emotions that are in various ways distorted. Perhaps they have an irrational hatred of members of a particular race or gender, or they are just systematically more sensitive to some things, perhaps taking them more personally than one ought to take them. One of the two, however, has come to recognize these facts about herself, and consciously corrects for them when forming beliefs, making decisions, etc. As long as this conscious correction is in any way a manifestation of her *being concerned for the good*, then I'd want to say that she is the more virtuous of the two people—even though, as I've stipulated, their emotions are equally off the mark. So unfortunately, it would be too simple to say that the degree to which a person's emotions are sensitive to objective moral values will always *match* the degree to which she is excellent in her concern for the good. But I think we *can* say—as I do in the argument below—that, *other things being equal*, if a person is virtuous, her emotions will be sensitive to objective moral values.

The second observation that I want to make about the MVP's emotional response to S's ϕ -ing goes back to my earlier reflections upon the nature of perception and its contribution to the justificatory status of our ordinary perceptual beliefs. As I observed above, ordinary perceptual beliefs are formed without deliberation, but still on a basis (perception) that is able to contribute positively to their justification. And this ability is rooted in the nature of perception—specifically, in its sensitivity (under certain conditions) to features of reality.¹⁴ Applying these points to the MVP's emotional response to S's ϕ -ing, the thing to observe is this: If the MVP forms the intuition that S's ϕ -ing was morally wrong on the basis of his anger response, then, other things being equal, the MVP's intuition is justified, even if it is formed on the basis of the anger response *alone* (i.e., and not also on the basis of any kind of deliberation, for instance, about his reasons for being angry, etc.). As we just saw, under certain conditions—namely, conditions of virtue—a person's emotions can be sensitive to moral values, just

¹⁴ This ability of perceptual experiences to contribute positively to the justification of perceptual beliefs depends upon more than *just* a sensitivity (under certain conditions) to features of reality. There is also the matter of there being some kind of relation between the conceptual contents of the perceptions and the subsequent contents of the beliefs. And this is a point on which there is apparently a difference (and perhaps, an important one) between ordinary perceptual beliefs and moral intuitions. In the case of the former, the conceptual contents of the experiences in which the beliefs are grounded presumably plays some role in shaping the contents of the beliefs themselves—I *believe* there's a red cup because I *perceive* a red cup. But this does not seem to be the case with respect to moral intuitions and the emotions in which they are grounded. The content of my intuition that Tom's action was morally wrong is presumably a *proposition* to the effect *that Tom's action was morally wrong*. But no part of the experience in which this belief is grounded—namely, my anger at Tom—has as part of its content the *wrongness* of Tom's action. And without a story about how certain emotional experiences relate to certain moral belief-contents, there may be a question about how the emotions can justify those beliefs. The good news is that I do think such a story can be told. Blair, for instance, suggests that, at some early stage in our evolution, humans engaged in a very primitive sort of “good/bad” judgment, and that this judgment was probably causally grounded in our affective responses to things—i.e., everything from natural disasters, to diseases, to the actions of others. Emotions with negative valences (anger, disgust, fear) gave rise to “bad”-judgments, and emotions with positive valences gave rise to “good”-judgments. If one couples this suggestion with the (plausible, I think) assumption that, along the way, our evaluative vocabulary (and the contexts in which we employed it) became more and more sophisticated, I think one will have a plausible account of the relation between emotions and the contents of moral beliefs up and running. The bad news, however, is that developing this story more fully goes beyond what I am up to in this dissertation, and so, will have to be put off until a later time.

as, under certain conditions, a person's perceptual faculties can be sensitive to other parts of the world. And if the sensitivity of perception to those parts of the world is such as to contribute positively to the justification of ordinary perceptual beliefs, then the sensitivity of emotion to moral values ought to make a similarly positive contribution to the justification of ordinary moral intuitions.

Here, then, is an argument for the (potential) trustworthiness of moral intuitions grounded in emotion:

1. Emotions are complex mental states consisting of construals, concerns, and feelings, related to each other in the way that I describe in chapter II.
2. *Psychological sentimentalism*: The psychological process ordinarily responsible for the production of moral intuitions—i.e., moral judgment—is one in which emotion figures centrally.
3. If (1), then (other things being equal) if I am virtuous, my emotions are sensitive to objective moral values.¹⁵

This is true if concerns figure into the nature of emotion in the way that I describe in chapter II, and also if moral virtue is a matter of aligning one's concerns with the good, as I say above in §1. My discussion of the MVP is meant to support premise 3; but importantly, premise 3 does not *depend* upon the MVP in any way. Rather, the MVP is meant only as a kind of image of how the account runs in the case of someone whose moral perceptual faculties are working as they should.

4. So, (other things being equal) if I am virtuous, my emotions are sensitive to objective moral values. (From 1 and 3)
5. If my emotions are sensitive to objective moral values, then they contribute positively to the justification of moral intuitions formed on their basis.
6. If my emotions contribute positively to the justification of moral intuitions formed on their basis, then, if my moral intuitions are formed on their basis,

¹⁵ I'll have more to say later about my use of the term 'sensitive' here.

then my moral intuitions are trustworthy.

7. So, (other things being equal) if I am virtuous, my moral intuitions are trustworthy. (From 2, 4–6)

Essentially, what I'm arguing is that emotions can amount to instances of what might be called moral perception. This in itself is not a new idea. In his *On Emotional Presentation* (1917), for instance, Alexius Meinong extends his broader theory of the intentionality of mental states to emotions in particular, arguing that (objective) moral and aesthetic values are presented to the mind via the emotions in essentially the same way that other features of reality are presented to the mind via what he calls ideas. He writes,

When I say, "The sky is blue," and then say, "The sky is beautiful," a property is attributed to the sky in either case. In the second case a feeling participates in the apprehension of the property, as, in the first case, an idea does.¹⁶

More recently, Lawrence Blum has argued that our perception of the morally salient features of one's circumstances is largely unreflective and emotional.¹⁷ And according to Maggie Little,

[I]n order to 'see' the moral landscape clearly, in order to discern it fully and properly, one must have certain desires and emotions. Caring, being outraged, being motivated to act—all these are part of discerning the moral features clearly. The ideal epistemic agent herself would have appropriate affect, for it is needed if one is to discern all that there is to see.¹⁸

What has been missing from these prior discussions, however, is an adequate account of just *how it is* that our emotions might come to be sensitive to moral reality in a way

¹⁶ Meinong 1972: 28-29.

¹⁷ Blum 1994.

¹⁸ Little 1995: 127.

similar to that of perception. It is one thing to suppose that we perceive moral reality by means of the emotions; it is another thing to explain how this happens. (Recall Wedgwood's complaint from chapter I.) I take this latter task to be necessary if the trustworthiness of our moral intuitions is to be secured; and so, this is the task that I have taken up here.

3. Roberts on emotion as an epistemic aid to moral judgment

During his recent tenure as the Alvin Plantinga Fellow at Notre Dame's Center for Philosophy of Religion, Robert Roberts spent some time extending his theory of emotion to a view of the role of emotion in moral judgment. The result is a view according to which "an emotion may be the perceptual basis of [...] moral judgment."¹⁹ This, of course, sounds very similar to what I've argued in this chapter, so it may be useful for me to take a moment to distinguish my own view from the one that he defends.

For illustrative purposes, Roberts focuses on a case from Dostoevsky's *Notes from Underground*, re-telling the story as if we know both the Underground Man and Liza:

Imagine that someone tells you the following story. A man you know personally has been insulted and rejected by his associates at a dinner party, and afterwards follows them to a brothel to start a fight, only to find that they have already dispersed into the rooms of the brothel. While there he falls in with a prostitute, Liza, whom you also know well enough to be concerned about her wellbeing. The man wants to assuage his wounded vanity, and has been in the habit of doing so by exercising power over others. After he has slept with Liza he preaches a little sermon to her on the glories of family life and the degradations of prostitution. He pours it on really thick, and by his rhetoric reduces her to a condition of bitter remorse, and of gratitude and admiration toward himself. Overplaying his assumed role of judge and savior he gives her his address on departing, indicating that she may come to him. When she does come to his room

¹⁹ Roberts unpublished.

several days later in hopes of pursuing the relationship with her sage redeemer, he is humiliated by her seeing his poverty and turns on her with vindictive anger, telling her that he never cared for her at all, doesn't mind if she degrades herself in prostitution, and was only using her to salve the social wounds he had received at the drinking party. In her disillusionment she is devastated and leaves.²⁰

As Roberts recognizes, our response to the Underground Man is likely to be something in the neighborhood of anger or indignation. And since our two views of the nature of emotion are similar in certain respects, I can agree with Roberts when he notes that these emotional responses are “based on a *concern* for Liza's wellbeing and a more general concern for justice.”²¹ But when he turns to explain how this emotional response might serve as a kind of epistemic aid to our moral beliefs about the Underground Man and his actions, there are two points upon which Roberts and I disagree.

First of all, Roberts does not distinguish—as I think he should—between our *emotional response* to the Underground Man and his actions, on the one hand, and our *moral evaluation* of him, on the other. As Roberts puts it, “In your experience of being angry at the Underground Man for his treatment of Liza, he appears to you as **culpable, bad, and deserving of hurt** for what he has done.”²² To be clear, Roberts thinks that we *construe* the man as culpable, just as one might construe the duck-rabbit figure as a duck (see chapter II, §4). And since Roberts and I are in agreement that construals are (sometimes) parts of emotion, he is apparently saying that the attribution of culpability here is not an *effect of* our anger response to the man, but rather *a part of* our anger response.

²⁰ Roberts unpublished.

²¹ Roberts unpublished; italics mine.

²² Roberts unpublished; bold in original.

Of course, one likely *will* think the man culpable, bad, and deserving of hurt; but whereas Roberts thinks that these are a *part of* our emotional response to the man and his actions, I would say that these thoughts are better understood as the *effects of* our emotional response—we think him culpable because of our anger with him.²³ For a number of reasons, I think it is important to keep *emotional response* and *moral evaluation* separate. For one thing, as I explained in chapter II, there seem to be clear cases in which one might have an emotional response without actually evaluating the object of the emotion in the relevant way, and vice versa. And for another thing, if we fail to sufficiently distinguish between emotional response and moral evaluation, it will be difficult to see how the former could be an *epistemic aid* to the latter, as Roberts thinks it can.

Second, apparently, Roberts only thinks that emotion can offer a kind of *ex post facto* epistemic support to moral beliefs. He begins by discussing three ways in which perceptual experiences can be epistemic aids. First, perception is an aid to *understanding*—i.e., we understand things better after we’ve perceived them. For instance, I might describe a painting to you in quite vivid detail; but once you see it for yourself, you’ll better understand the painting that I described. Second, perception can *augment justification*. You may tell me that there’s a deer walking through the backyard, and on the basis of your testimony, I may be justified in believing this. But when I look for myself, my belief is more justified than it was prior to my perceptual experience of

²³ One might worry that there’s a “chicken-or-the-egg” problem here: do we think him culpable because of our anger, or are we angry because we think him culpable? In light of the evidence surveyed in chapter III, I think the causal sequence here begins with the anger response. But I also recognize that this unconscious causal process often happens so quickly as to give the appearance of either co-occurrence (of emotion and evaluation) or even a reverse-order sequence.

the deer. And finally, even if perception doesn't augment justification, Roberts thinks that the *personal acquaintance* provided by perception is still a kind of epistemic upgrade. Suppose a bird expert tells me that there's a rare bird in the tree outside my office. The combination of his expertise and my total ignorance of rare bird species may conspire to make it the case that my perceptual experience of the bird does nothing at all to the justification of my belief that there's a rare bird in the tree. But still, Roberts thinks, in virtue of becoming personally acquainted with the bird in the tree (by seeing it for myself), I upgrade my epistemic situation with respect to my beliefs about the bird.

There's nothing quite like perceiving for yourself, whether what's perceived is the flavor of a fine whiskey, the nastiness of a case of racial injustice, the gular area of a cormorant, the necessity of a necessary proposition [...] The upgrade here seems to be a matter of epistemic proximity or intimacy with the object.²⁴

These are the ways in which perception can be an epistemic aid, according to Roberts. And it is interesting to note that they are *all* cases of *upgrades* to epistemic situations—i.e., they are all cases in which a subject's *current* epistemic situation with respect to some belief or set of beliefs is *made better* by perception. In the same way, says Roberts, emotional responses can upgrade the epistemic status of our moral beliefs: We *understand* the injustice of the Underground Man's behavior better after we've felt it; we're more justified in believing that the behavior is unjust after we've felt it; and the sort of personal acquaintance with injustice that the emotions provide can itself be a kind of epistemic upgrade.

While Roberts may be right about all of this, these are not the only, or even the main, sorts of epistemic support that perceptual experiences offer to perceptual beliefs.

²⁴ Roberts unpublished.

The primary sort of support that perception offers to perceptual beliefs is a causal one: our perceptual beliefs are *prima facie* justified because they were *caused by* perceptual experiences that themselves reliably track the truths of perceptual beliefs. My belief that there is a deer in the backyard would be more than sufficiently justified *without* your testimony, on the basis of my perceptual experience *alone*, because (other things being equal) the perceptual experience itself arises out of a process that reliably tracks truths of the relevant sort. In other words, perceptual experiences are not only good for *upgrading* epistemic situations (i.e., by augmenting a belief's justification); they are equally capable of getting perceptual beliefs off on solid epistemic footing in the first place. Likewise, emotion may be able to upgrade the epistemic status of certain moral beliefs, in precisely the ways that Roberts explains. But on the account that I defend in this chapter, emotion, like perception, is able to get ordinary moral intuitions off on solid epistemic footing in the first place.

Furthermore, Roberts never makes it clear *why* it should be the case that emotion can be *any* kind of epistemic aid to moral judgment. Again, he rightly observes that our emotional responses to the Underground Man are rooted in concerns for such things as justice and Liza's wellbeing. But from there, he ends up just *stating* that our anger or indignation here provides some epistemic support to our beliefs about the blameworthiness of the man, or the injustice of his behavior. At no point does Roberts acknowledge that it is *because* of the concerns in which our emotional responses are rooted—and what this indicates about the state of our moral character—that the emotional responses are, in this case, epistemic aids. If we were similarly angry or indignant in response to the Underground Man, but the anger or indignation were rooted

in concerns that are totally irrelevant to moral goodness, then our emotional responses surely would not offer any kind of epistemic support to a belief in the injustice of the man's behavior. So while I agree with Roberts in thinking that emotions can be epistemic aids to moral beliefs, I think an account of the relation between emotion and moral virtue—of the sort that I describe in §1—is needed in order to explain why this is the case.

4. Reliabilism and moral reliabilism

Those familiar with theories of justification and knowledge will perhaps recognize the view that I defend in this chapter as a form of process reliabilism about justification. Roughly speaking, process reliabilism is the view that, other things being equal, beliefs are epistemically justified if they result from a belief-forming process that is reliable, where reliability is typically understood in terms of conduciveness to truth. If belief-forming process *P* produces true beliefs at some sufficient (though hard to specify) frequency, then *P* is reliable; and if *P* is reliable, and belief *b* is a product of *P*, then *b* is justified. According to the view that I defend here, moral judgment, sufficiently guided by moral virtue, is reliable; and so, the moral intuitions of the sufficiently virtuous are justified.

Now, I've been reluctant (until now) to ally my view too closely with reliabilism, because I do not want to be responsible for defending process reliabilism *in general* in order to adequately defend my view of moral judgment. Limitations of time and space simply will not allow such a thing—and besides, it has already been done much more

clearly and forcefully than I'd be able to do here.²⁵ But it may nonetheless be helpful, if only for the purposes of clarifying my overall view, for me to say a few things about how my own brand of moral reliabilism might offer responses to a few objections often leveled against process reliabilism in general.

One objection that is often made to reliabilism is that reliability is not *necessary* for a belief to be justified. This is typically demonstrated by means of a thought experiment in which people in one world have perceptual beliefs that are qualitatively identical to those of people in the actual world, but are nonetheless brought about by a *non-reliable* process—e.g., manipulation by an evil demon. The perceptual beliefs of those in the non-actual world are surely justified—after all, they're identical to our perceptual beliefs, and ours are justified—but, by stipulation, they were not reliably formed. So reliability is not necessary for justification.

If I intended to be offering a complete analysis of moral justification, and doing so in terms of reliability, then this objection might be a problem for me. But that is *not* what I'm doing in this dissertation. I am not saying that moral intuitions (or moral beliefs, more generally) could *only* be justified by being reliably formed. Remember, I am responding to skeptics who think that moral intuitions being grounded in emotion is a reason to think that they are not trustworthy (and so, not justified). I reject this. I think that moral intuitions can be justified *because* are grounded in emotion, and I explain how in this chapter. This is perfectly compatible with there being a whole host of other ways—besides being reliably formed—for moral intuitions to be justified.

²⁵ See Goldman 2012.

More interesting, however, is another objection to reliabilism according to which reliability is not *sufficient* for justification. Indeed, I do want to make the sufficiency claim, so I will have to say something more in response to this objection. As the objection typically goes, we can imagine a case in which a subject has a perfectly reliable faculty for forming certain beliefs—where, again, ‘reliable’ is meant in the sense of being conducive to truth, or to the formation of true beliefs—but either has no reason to believe that it is reliable, or even has good reason for doubting that it is reliable. In such cases, some think that the subject would not be justified in holding whatever beliefs result from this process, reliable though it may be. So reliability is not sufficient for justification. And in the case of moral intuitions, it is indeed quite easy to imagine someone doubting that a psychological process in which emotion figures centrally could be reliable for the production of true moral beliefs. As I have been saying from the start, this is apparently a very popular opinion among both philosophers and non-philosophers alike. If I take myself to have good reason to doubt the reliability of my own moral judgment upon learning that it is a process in which emotion figures centrally—maybe I’ve noticed what sorts of distorting influences my emotions can have upon my thinking and deliberation about moral matters; or maybe I agree that emotion is often shaped by virtue, but I doubt that I am virtuous—how could I then be justified in holding whatever beliefs it happens to produce (even if it is, in fact, reliable)?

I have two things to say in response to this worry. First, I think a hardline reliabilist will not—and should not—be bothered by cases like this. The fact of the matter is that we are *often* wrong about the justificatory status of our own beliefs. Of course, it is probably more common for us to err on the side of *overestimating* how

justified a particular belief really is, but it should come as no surprise that we sometimes underestimate our own justifiedness as well. If a moral intuition of mine was reliably formed, then any doubts about its justification are *false*, even if they are had for good reason. I may not *know* that I am justified until I am able to remove these doubts; and I may struggle to justify the intuition *to others*; but for all that, the intuition itself is justified, in virtue of being reliably formed. So the first thing to say in response to the worry about reliability not being sufficient for justification is just that I endorse this hardline reliabilist response.

The second thing to say is aimed more specifically at the worry *as it applies to emotions and moral intuitions*. There may indeed be good reasons for one to doubt the extent to which one's emotional responses track moral values, or the degree to which one is morally virtuous. But my suspicion is that there is only very little overlap between the set of people who have serious doubts about these things and the set of people who *ought* to have serious doubts about these things; and I think the latter set is much, much smaller than one might suspect. Now, it would be hard for me to substantiate these suspicions empirically, because, as far as I can tell, there just is not a satisfactory method for measuring such things as the degree to which a person is morally virtuous. But I nonetheless take myself to have given reasons in this chapter for thinking that worries about our emotions and their ability to generate trustworthy moral intuitions are often overblown. We may not be *maximally* virtuous, but maximal virtue is not necessary for moral intuitions to be justified—just as a perceptual faculty would not have to function perfectly in order for perceptual beliefs to be justified. I'll have more to say about this in the next section, and still more in the next chapter.

There is another problem that may confront versions of *moral* reliabilism in particular, and it has to do with the causal relationship (or lack thereof) between moral beliefs and moral values. In defending the trustworthiness of moral intuitions, I drew upon an analogy to the trustworthiness of ordinary perceptual beliefs. But, some might argue, there is an important difference between these two types of beliefs. In the case of perceptual beliefs, it is common to think that they are reliably formed because of the causal relationship that exists between perceptual experiences and the things perceived. The sweet sensation is *caused by* the tea, for instance, so of course a belief that the tea is sweet is reliably formed if it is formed on the basis of the sweet sensation. But if this is what it takes for our fast and unreflective practice of forming perceptual beliefs to count as reliable, then if one wants to say of our ordinary moral intuitions that they, too, are reliably formed, it would seem one will have to adopt a rather hearty—not to mention extremely controversial—metaphysics of morals according to which moral facts or properties are able to somehow cause our emotional responses. I’m trying to give an account of moral judgment according to which its products are (or can be) justified beliefs about objective moral values, but I also want to remain as neutral as possible about the metaphysics of those values.

Thankfully, as has been pointed out by Russ Shafer-Landau—himself a defender of moral reliabilism—causation is not necessary for reliability.²⁶ That is, beliefs about *x* can be reliably formed even if *x* itself does not enter causally into the production of those

²⁶ Shafer-Landau 2003: 275. Another problem that often confronts versions of process reliabilism is the so-called Generality Problem. I ignore the Generality Problem here; but for a helpful discussion of the problem, and one that confronts the problem in the context of moral belief and justification, see Shafer-Landau 2003: 280-285.

beliefs. One way to establish that this is the case would be to tell a story about how beliefs about x (or the process by which these beliefs are formed) *track* the relevant x -facts. If there is a tracking relation of some sort, then the beliefs might be reliably formed even if they are not caused by x . So if our moral intuitions are formed on the basis of our emotional responses to things, they may yet be justified even if moral values do not themselves play a role in causing the emotions—for it might nonetheless be the case that our emotions *track* moral facts or properties to a sufficient (though hard to specify) degree. Of course, we should want for it not to be an *accident* that our emotional responses track moral values in this way. But the account I’ve given of the nature of emotion and its relation to moral virtue is meant to ensure that this is not the case.²⁷

5. Three lingering worries

Let me briefly address a few lingering worries about my view before returning in the next, and final, chapter to the skeptical arguments of Greene, Street, Nichols, and Sinnott-Armstrong. The first of these worries has to do with the *objectivity* of moral judgment. I’ve attempted in this chapter to give an account of moral judgment according to which the products of moral judgment are (or can be) justified beliefs about objective moral values. But, it is natural to wonder, if moral judgment ordinarily goes the way of the emotions, and my emotions are to a significant degree determined by *my concerns*, then how could the products of moral judgment be anything *but* subjective? And, one might add, it is not enough for me to say that our emotions are sensitive to an objective moral reality as long as they are rooted in a concern *for the good*, for conceptions of the good

²⁷ Looking back at the argument in §2, we might understand the term ‘sensitive’ in premises 3-5 in terms of kind of *tracking* relation of the sort described here—i.e., the effect of virtue upon moral judgment is that of ensuring that emotional responses *track* moral values, even if they are not *caused*, in any sense, by moral values. It is in this way that our emotions can be *sensitive* to moral values.

are likely every bit as varied and subjective as anything else.

As I mentioned in the introduction of this chapter, one way to avoid this worry might be to conceive objective moral values as secondary, or response-dependent, qualities. If being able to cause negative emotions (under certain circumstances) is *all it is* for ϕ -ing to be morally wrong, for instance, then it should be pretty easy for my moral intuitions about the wrongness (or not) of ϕ -ing to count as beliefs about objective moral values, even if they are grounded in nothing *but* emotion. But as I noted earlier, there may be reasons for preferring an account that does not depend upon a secondary quality view of moral values.

In response to this worry about objectivity, then, I'll say this: when I claim that moral virtue consists in being excellently concerned for the good, what I mean to say is that virtue consists not in being concerned for *what one takes to be* the good (or even what one *would* take to be the good under ideal conditions), but rather in being concerned for *what is in reality* the good. So for instance, if a utilitarian exhibits excellence in his concern for maximizing happiness, he will only count as virtuous on this view if it turns out that goodness and the maximization of happiness are one and the same. On the assumption that there are objective moral values in the world, moral virtue consists in being excellently concerned for what is objectively *the good*. And to the extent that someone's concerns are aligned with *it* (regardless of whether or not she is aware of this alignment), I contend, she will be disposed to respond emotionally to objective moral values.

A second worry stems from the obvious and unfortunate fact that we are not maximally virtuous. It might be easy to concede that someone like the MVP is attuned to morality in the way that I describe, but what about the rest of us? Should the disparity between the MVP and ordinary people like us not inspire some doubt as to the trustworthiness of our moral intuitions?

Now, in one sense, it simply does not matter that we are not maximally virtuous. Skeptics like Nichols and Sinnott-Armstrong argue that moral intuitions are probably not trustworthy if they are grounded in emotion—that is, there is supposed to be something *about emotion* that is incompatible with the production of trustworthy moral intuitions. The MVP is a counter-example to their arguments: his moral intuitions are trustworthy *because* they are grounded in emotion. What these skeptics will have to add to their arguments, then, is a further argument to the effect that the disparity between the MVP and the rest of us is *large enough* to render *our* moral intuitions untrustworthy. But I should say that it is hard to even imagine what such an argument would require. These skeptics would presumably need to support their disparity claim with empirical evidence, but as far as I can tell, none of the evidence currently on offer is adequate to the task. There is, of course, plenty of evidence—empirical and anecdotal—for thinking that our emotions sometimes cloud moral judgment. But then, there is also plenty of evidence for thinking that our *perceptions* sometimes cloud perceptual judgment—think, for instance, of cases of illusion, hallucination, and the like. If our ordinary perceptual beliefs are no less trustworthy because of this, why think any differently of our moral intuitions? (In the next chapter, I discuss cases of so-called ‘moral illusion’.)

Recall the analogy that I drew between the MVP and someone with a “musical ear.” It is surely true that my wife exhibits a greater sensitivity to features of music and musical performances than those of us who are not as highly trained. But if you were simply to hammer away at any adjacent keys on a tuned piano, probably everyone within earshot would cringe. With respect to a great many breaches of melody, harmony, etc., it doesn’t take an expert musician to hear them. Likewise, with respect to many moral matters, it doesn’t take a maximally virtuous person to feel them. We all react with horror to stories of such things as child abuse, torture for pleasure, genocide, and other obvious moral wrongdoings. And apparently, in light of the evidence from chapter III, these emotional reactions play an important role in causing us to believe the things we do about such actions. We know the answers to a lot of moral questions, and, I’d argue, we have our emotions to thank for this.

If there is a lesson to be learned from the fact that we are not maximally virtuous, I think it is primarily a *practical* one. Rather than systematically doubting the deliverances of our moral sense, we should simply proceed with caution, and do what we can to sharpen our sense of the good.²⁸ It may be perfectly reasonable to recognize in oneself tendencies for specific concerns to be otherwise than they should—e.g., perhaps it’s been brought to your attention that you have an insufficient degree of self-respect. And consequently, one may reasonably doubt the moral intuitions that one has related to these misplaced or misguided concerns—e.g., you begin to doubt intuitions to the effect that it is morally okay for your boss to treat you in this or that way. But this doubt will,

²⁸ Of course, selecting a strategy for doing so will likely require us to regard some moral intuitions as incorrigible. As far as I can tell, this is not a problem for the point that I am making here.

or should, always be localized to the area of concern, and should not be cast over all of our moral intuitions—unless one somehow recognizes in oneself serious and widespread defects in either emotion or moral character.

Finally, rather than worrying that we are not virtuous *enough* for our moral intuitions to be trustworthy, one might wonder if anyone is virtuous *at all*. Here I have in mind another area of research in empirical moral psychology—i.e., research into the psychological reality of *traits* (e.g., traits of virtue and vice). According to some so-called *situationists*, there really is no such thing as moral virtue, if by ‘virtue’ we mean to denote a particular class of traits. And this is because, they argue, there is insufficient empirical evidence to support our ordinary, everyday attribution of traits to people. As Gilbert Harman puts it, “Empirical studies designed to test whether people behave differently in ways that might reflect their having different character traits have failed to find relevant differences.”²⁹ What *do* seem to make noteworthy differences to a person’s behavior are features of his situation that often seem to have little or nothing to do with virtue—e.g., people are apparently significantly more likely to engage in helping behavior if they’ve just found a dime in the coin-return slot of a pay phone.³⁰ Thus, Harman continues, “ordinary attributions of character-traits to people may be deeply misguided, and it may even be the case that there is no such thing as character.”³¹ And of course, if there is no such thing as character, then the view of moral judgment that I defend in this chapter will fall flat on its face, since there will be nothing to ensure that

²⁹ Harman 1999: 316. See also Doris 2002.

³⁰ Isen and Levin 1972.

³¹ Harman 1999: 316.

our emotional responses are tracking moral truths in the sort of justification-conferring way that I describe.

As interesting as I find the literature on the psychology of moral character, a full response to the situationist would take me too far away from the focus of this chapter, which is the trustworthiness of moral intuitions. Suffice it to say, for now: situationists typically conceive of traits of virtue as direct dispositions to particular behaviors, and I agree with those who argue that this is a very problematic conception of virtue and moral character.³² (Notice that the view of moral virtue that I take up in §1 does not obviously commit me to *anything whatsoever* with respect to the connection between virtue and behavior. To say that a morally virtuous person is someone who is excellent in her concern for the good is to say nothing at all about how she'll behave in a given situation.) If something like this response is on point, then we may be free to grant that situationists have established the non-existence of *something*, just not moral virtue.

6. Concluding remarks

I began this dissertation by noting just how important moral intuitions are to us, both within and outside of philosophy. Then I explained how recent work in empirical moral psychology has raised what many think is a serious worry about the trustworthiness of our moral intuitions: the evidence suggests that these intuitions are ordinarily the effects of emotion, and emotion is typically thought to be something that could only *get in the way* of the production of trustworthy moral beliefs. The main goal of this dissertation, then, as I described it there, has been to defend the epistemic status of moral intuitions by

³² See, e.g., Kamtekar 2004 and Adams 2006 (especially ch. 8).

showing how it can be that they are trustworthy *when*, and *because*, they are grounded in emotion. That is what I have done in this chapter.

As readers can now clearly see, my case very much depends upon the nature of emotion being as I describe it in chapter II. For that account of the nature of emotion allows me to develop the account of the relationship between emotion and moral virtue that I do here in chapter IV. And as long as these two accounts are on point, it turns out that our ordinary moral intuitions can sometimes be trustworthy *because* they are grounded in emotion, and not *despite* this fact. And if that's correct, then apparently, we have good reason to go on as if our moral intuitions are trustworthy—even if we are unaware of this.

In the next and final chapter, I'll look back at the four skeptical arguments that I mentioned at the end of chapter III, and I'll show how the view that I've defended in this chapter provides me with decisive responses to these arguments. I'll also look at another skeptical argument, also from Sinnott-Armstrong, and explain why I think this argument actually provides further support for my views.

V
EMOTION AND MORAL JUDGMENT SKEPTICISM

In chapters II-IV, I developed an account of moral judgment according to which, first, emotion figures centrally as a cause in the production of moral intuitions, and second, moral intuitions are (or at least, they *can be*) trustworthy *because* they are grounded in emotion, and not despite this fact. The time has come for me to put the view to the test, and to show that it can deliver on its promise to rebut a certain strain of moral skepticism. For the purposes of this chapter, I will call anyone who thinks that moral intuitions cannot be trustworthy if they are grounded in emotion—or at least, that their being grounded in emotion is a reason to think that moral intuitions are not trustworthy—a *moral judgment skeptic*. (As a reminder: as I am using the term, moral intuitions are *trustworthy* if they really are justified beliefs about objective moral values.) So the questions left for me to address are these: What reasons are there for moral judgment skepticism? And are they reasons that my view can overcome?

In §§1-3, I examine four arguments for moral judgment skepticism, showing how my view can respond in each case. In §1, I look at an argument from Joshua Greene, and then a similar, but much stronger, argument from Sharon Street. Both arguments focus on the evolution of emotions, and how this might tell against thinking that emotions can reliably track objective moral values, as I argue they can. In §2, I turn to an argument from Shaun Nichols. Of concern to Nichols is the apparent irrationality of emotion.

Since emotions are “rationally arbitrary,” he thinks—and since emotion figures centrally in moral judgment (he, too, is a psychological sentimentalist)—moral judgment cannot be objective. And in §3, I discuss an argument from Walter Sinnott-Armstrong. Whereas Greene, Street, and Nichols are chiefly concerned with the *objectivity* of moral judgment—i.e., whether or not moral judgment is the sort of process that can be counted on to produce beliefs about objective moral values—Sinnott-Armstrong is concerned more with the *justification* of moral intuitions. Since emotions are in the business of clouding moral judgment, he thinks, we’ll need some sort of further inferential confirmation in order to justify our moral intuitions.

Then, in §4, I’ll briefly consider another sort of argument for thinking that moral intuitions are untrustworthy. This argument also comes from Sinnott-Armstrong, and it has to do with the apparent susceptibility of moral judgment to what might be called “moral illusions.” Rather than *denying* that moral judgment is susceptible in this way, though—in fact, I agree that it is—I’ll instead explain why (a) the susceptibility of moral judgment to such illusions might be seen as further confirmation of the view that I developed in chapters II-IV, and also (b) why this susceptibility is really not the problem that Sinnott-Armstrong thinks it is. And finally, in §5, I’ll conclude the dissertation by summarizing my overall view, and then speculating about some further implications that the view may have upon other issues in ethics and moral psychology.

1. Greene, Street, and the evolution of emotions

One argument for moral judgment skepticism comes from psychologist Joshua Greene. He sets the argument up by briefly describing the difference between realists and anti-

realists in metaethics, and then suggesting, optimistically, “I believe that neuroscience and related disciplines have the potential to shed light on these matters by helping us to understand our common-sense conceptions of morality.”¹ According to Greene, metaethicists are primarily concerned with the (metaphysical) nature of moral values, and there are basically just two sides that one might take on this issue: either one affirms that truths about the moral values of things are “full-blown truths, mind-independent facts about the nature of moral reality,” or one affirms that moral values are, “like sexiness, in the mind of the beholder.”² And he proposes that we attempt to resolve this issue by “examin[ing] what is in the minds of the relevant beholders,” i.e., by determining *just what is* going on in a person’s mind when she judges a thing to have this or that moral value. Understanding how moral intuitions are formed, he suggests, may help us to determine if they are best understood as beliefs about objective moral values or, instead, as “projections of internal attitudes.”³

At this point, Greene makes three observations about moral judgment. His first observation is that moral judgment ordinarily operates very quickly—even automatically—and effortlessly (see §3 of chapter I), and that this gives to moral judgment its distinctly perceptual phenomenology.

An interesting feature of many intuitive, effortless cognitive processes is that they are accompanied by a perceptual phenomenology. For example, humans can effortlessly determine whether a given face is male or female.⁴

¹ Greene 2003: 849.

² Greene 2003: 849.

³ Greene 2003: 849.

⁴ Greene 2003: 849.

His second observation is that it is *because* of its perceptual phenomenology that moral judgment *seems to us* like a process by which we form beliefs about mind-independent, objective moral values. After all, in ordinary cases, perception is a matter of being impressed upon by things that are *out there* in the world to be perceived. And finally, his third observation is that *emotion* is what underlies the perceptual phenomenology of moral judgment. Citing research in empirical moral psychology—like that discussed in chapter III—Greene explains that moral judgment is able to operate with such speed and effortlessness because it is primarily a matter of automatic emotional responses. In other words, psychological sentimentalism is true, and this explains why moral judgment has the perceptual phenomenology that it has.

So far, Greene has said nothing that I find particularly bothersome. I might argue that it is more than just the perceptual phenomenology of moral judgment that inclines us to believe that it is a process by which we form beliefs about objective moral values. But he is probably right that this phenomenology contributes to our going on as if moral judgment is objective. It is when he goes on to draw out a supposed implication of these three observations about moral judgment, however, that Greene and I disagree.

Essentially, Greene argues that, if it is indeed emotion that gives moral judgment its perceptual phenomenology, then it must be the case that its perceptual phenomenology *is a sham*. For this would imply that moral judgment is like aesthetic judgment, in that our beliefs about the moral values of things—like our beliefs about a thing’s “sexiness”—are grounded (causally) in emotion (or some other affective state). And, Greene simply assumes, everyone will agree that the perceptual phenomenology of

aesthetic judgment is a sham, since there are no “full-fledged truths” about such matters as beauty and sexiness.

We have here the beginnings of a debunking explanation of moral realism: we believe in moral realism because moral experience has a perceptual phenom-enology, and moral experience has a perceptual phenomenology because natural selection has outfitted us with mechanisms for making intuitive, emotion-based moral judgments, much as it has outfitted us with mechanisms for making intuitive, emotion-based judgments about who among us are the most suitable mates. [...] According to this view, moral realism is akin to naïve realism about sexiness, like making the understandable mistake of thinking that Tom Cruise is objectively sexier than his baboon counterparts.⁵

In other words, moral judgment is relevantly similar to aesthetic judgment, and aesthetic judgment is clearly not objective. So moral judgment must not be objective either, despite its perceptual feel.

As Greene presents it here, this is a pretty weak argument. For one thing, I think there are many who will find his remarks about aesthetic judgment rather controversial. Certainly, to some degree, we think that things like sexiness are “in the mind of the beholder.” But Greene assumes that we’ll agree that *all* aesthetic matters are like this. I suspect he is wrong about this—would people really agree that the superiority of a renowned artist’s work to an average person’s doodles is entirely in the mind of the beholder?—but I’ll not pursue the matter any further here. A question I would like to pursue, however, is this: Why not think that moral judgment is like gender perception? By Greene’s own lights, gender perception is *another* fast and effortless psychological process for which we have evolution to thank; and yet, surely he’d not say that the perceptual phenomenology of gender perception is a sham. After all, there certainly are mind-independent facts of the matter about whether a particular face is that of a male or

⁵ Greene 2003: 849.

female. Here is Greene's explanation of the difference, which he adds immediately after the above quotation:

Note that according to this view moral judgment is importantly different from gender perception. Both involve processes that give rise to a perceptual phenomenology, but in the case of gender perception the phenomenology is veridical: there really are mind-independent facts about who is male or female.⁶

But of course, no moral realist will accept this explanation, for it only begs the very question at hand. Greene raised a *psychological* question about what is going on in the minds of moral believers, in the hope that it would point us in the direction of an answer to a *metaethical* question about the nature of moral values. But in the end, it looks like he moves in precisely the opposite direction: he *assumes* an answer to the metaethical question—i.e., he assumes that there aren't any mind-independent moral values—and then, on the basis of this assumption, he gives us an answer to the psychological question: *since* there are no objective moral values for us to perceive, it cannot be the case that genuine (moral) perception is going on in the minds of moral believers.

Greene will likely respond that I am failing to appreciate what its *emotional* nature implies about the objectivity of moral judgment: moral intuitions, unlike beliefs about gender (presumably), are grounded in emotion. And *that's* why it cannot be the case that moral intuitions are beliefs about objective moral values. But then Greene would only be begging the question against those, like myself, who argue that emotions *can* be instances of moral perception. As I explained at the end of chapter III, Greene apparently shares the assumption that there is something *about emotion* that makes it the case that moral judgment cannot be objective if emotion plays a significant role in the

⁶ Greene 2003: 849.

production of moral intuitions. But he does not offer us any reason to share this assumption *with him*—he just assumes we will—and in chapter IV, I provide grounds for rejecting it.

A more interesting argument—and one similar to Greene’s—comes from Sharon Street.⁷ At the core of Greene’s argument is the suspicion, as he puts it, that

[u]nderstanding where our moral instincts come from and how they work [...] can lead us to doubt that our moral convictions stem from perceptions of moral truth rather than projections of internal attitudes.⁸

Street shares this suspicion. Reflecting upon the evolutionary origins of human moral judgment, she thinks, will lead us to doubt that our minds are ever in the business of apprehending objective moral values.

Street’s argument begins with the observation that the forces of natural selection likely had a tremendous influence upon the shaping of human attitudes and emotions, and consequently, upon the shaping of what she calls our “evaluative tendencies.” Surely, for instance, evolution is at least largely responsible for shaping our emotional responses to such things as death, suffering, and survival. Consequently, humans in general treat the fact that something would promote one’s survival *as a reason in favor of it*—and, here again, we have evolution to thank for this. I imagine few would disagree with Street on this point. However, on the basis of this observation about the tremendous influence of natural selection upon human emotions, Street goes on to raise a dilemma for anyone,

⁷ It may be worth noting that, despite the similarities between their two arguments, Greene and Street are up to significantly different things. For one thing, while Greene is engaged only in a discussion of psychological sentimentalism, Street takes herself to be engaged in a project that is much broader than this—namely, that of using Darwinian evolutionary considerations to argue against realism in metaethics. So when I ask, in a few paragraphs, how a psychological sentimentalist might respond to Street’s argument, I do *not* mean to imply that Street’s argument is aimed at psychological sentimentalists in particular.

⁸ Greene 2003: 850.

like myself, who thinks that there are objective moral values for our minds to apprehend when engaging in moral judgment.

Briefly, the dilemma is this. We must either *affirm* or *deny* that there is a relation of some sort between objective moral values and the forces of natural selection. If we affirm that there is a relation, Street thinks, it could only be a tracking relation: evolution outfitted us with emotions and attitudes that reliably track the moral status of things.⁹ The problem with this, however, is simply that there are *better* explanations (understood in terms of scientific adequacy) of the evaluative tendencies of human beings than the one suggested by the tracking account. Take, for instance, Street's own adaptive link account, according to which

tendencies to make certain kinds of evaluative judgments rather than others contributed to our ancestors' reproductive success not because they constituted perceptions of independent evaluative truths, but rather because they forged adaptive links between our ancestors' circumstances and their responses to those circumstances, getting them to act, feel, and believe in ways that turned out to be reproductively advantageous.¹⁰

⁹ Why think that it could *only* be a *tracking* relation? As Street explains, it is because of the very nature of moral realism. "[B]ecause it views these evaluative truths as ultimately independent of our evaluative attitudes, the only way for realism *both* to accept that those attitudes have been deeply influenced by evolutionary causes *and* to avoid seeing these causes as distorting is for it to claim that these causes actually in some way tracked the alleged independent truths. There is no other way to go" (Street 2006: 134-135).

¹⁰ Street 2006: 127. Street apparently uses the term 'judgment' to refer either to a *propositional attitude* of some sort (e.g., a belief), or to the *act of forming* such an attitude (see §3 of chapter 1). One reason that I try *not* to use 'judgment' in this sense, however, is that there may be different types of moral, or evaluative, *beliefs*, and these different types of beliefs may be the products of different types of belief-forming *acts*. Following Sidgwick, for instance, Shafer-Landau distinguishes between three different types of moral beliefs: (i) beliefs expressing general moral principles (e.g., the belief that one should act only on maxims that one could will to be universal laws, etc.), (ii) beliefs expressing moral appraisals of act-types (e.g., the belief that lying is wrong), and (iii) beliefs expressing moral appraisals of act-tokens (e.g., the belief that it was not wrong for Suzy to tell lie *L* in circumstance *C*). These strike me as importantly different kinds of beliefs, and it would not be surprising if they resulted from very different kinds of belief-forming acts, or psychological processes. Consequently, to use 'moral (or evaluative) judgment' to refer to *any* evaluative belief, or to *any* act by which we form such beliefs, as Street apparently does, one risks saying things that are true only in some cases, but false in others. For instance, while it may seem quite plausible that we have evolution to thank for the mechanisms by which we form beliefs about act-tokens, is it equally plausible that beliefs expressing general moral principles were influenced in the same way, to the

As Street goes on to argue (I won't recall the details here), the adaptive link account is clearer, more parsimonious, and explanatorily more powerful than the tracking account. So, even though Street thinks it more promising for the realist to affirm, rather than deny, a relation between the forces of natural selection and objective moral values, this nonetheless turns out to be quite a tough row to hoe.

Why not *deny* a relation between evolutionary forces and objective moral values? According to Street, the problem here is that there is no way for the realist to avoid the conclusion that it would then be an unbelievable—literally, *unbelievable*—coincidence if any of our ordinary moral intuitions were actually true.

On this view, allowing our evaluative judgments to be shaped by evolutionary influences is analogous to setting out for Bermuda and letting the course of your boat be determined by the wind and tides: just as the push of the wind and tides on your boat has nothing to do with where you want to go, so the historical push of natural selection on the content of our evaluative judgments has nothing to do with evaluative truth.¹¹

And it is no help to insist that, at some point in our evolutionary past, we developed the capacity to use rational reflection as a means to “right the ship,” as it were—i.e., as a means to counter the distorting influences of evolution, and steer moral judgment in the direction of objective moral values. For when we engage in rational reflection, Street explains, it is not as if we *step outside of* our own evaluative standpoint, detached entirely from our own evaluative beliefs. Rather, the person engaged in rational reflection “*uses* them [evaluative beliefs], reasons in terms of them, [and] holds some if them up for

same degree, etc., by the forces of natural selection? I think probably not, though I'll not pursue the matter any further.

So that it fits more neatly into the discussion that I've been carrying on in this dissertation, I might recast Street's argument as one that concerns the influence of evolutionary forces upon moral judgment, i.e., upon our tendency to form certain moral intuitions, and not others.

¹¹ Street 2006: 121.

examination in light of others.”¹² And of course, these are the same beliefs that are supposed to have been shaped to a significant degree by the distorting forces of natural selection.

So long as we assume that there is no relation between evolutionary influences and evaluative truth, the appeal to rational reflection offers no escape from the conclusion that, in the absence of an incredible coincidence, most of our evaluative judgments are likely to be false.¹³

So if we deny a relation between the forces of natural selection and objective moral values, then it seems we’re quickly led to the unfortunate conclusion that most of our ordinary intuitions about the moral status of things are probably false, since they are grounded causally in emotions that have been shaped by evolution in ways that have nothing whatsoever to do with objective moral values.

How might a psychological sentimentalist who *also* wants to affirm the existence of objective moral values escape Street’s “Darwinian dilemma”? Well, as I’ve mentioned before, one could adopt a secondary quality view of moral values, like the one that McDowell endorses. If *being such as to elicit certain emotional responses* is all it is for *x* to be morally wrong, and this fact about *x* obtains sufficiently independently of our thoughts about its capacity to elicit those responses, etc., then one could apparently claim, perfectly consistently, that (a) *x* is objectively morally wrong, (b) our intuitions about the wrongness of *x* are (sometimes) causally grounded in our emotional response to *x*, and (c) evolutionary forces have had a tremendous influence upon our emotional responses to things. But what if one would rather *not* adopt a secondary quality view of moral values?

¹² Street 2006: 124; italics in original.

¹³ Street 2006: 124-125.

Think back to my account of the nature of emotion, and ask: *How* is it that evolution has had the sort of influence that it has apparently had upon our emotions? I'd answer that this influence is a result of the ways in which evolution shaped our *concerns* and *construals*. For instance, a concern for our own survival is among the deepest and most universal of human concerns, and there is no doubt that the forces of natural selection had something to do with this. Evolution is also (partly) responsible for shaping our construals of ferocious animals *as threats*, of some other people *as physically attractive*, of children *as vulnerable*, and so forth. (Indeed, this may be a consequence of evolution's prior effects upon our *concerns*.) In these ways, my account of the nature of emotion helps to explain how it is that evolution could have shaped our emotional responses to things.

My account of emotion can also help to explain how it is that moral intuitions grounded in emotion can be trustworthy *despite* the “purely distorting” influences of evolution, especially when it is coupled with the account of moral virtue described in the last chapter. Notice that nothing Street says is incompatible with it *now* being the case that many of us are morally virtuous in this way—i.e., concerned for things that are objectively morally good, in the right ways, to the right degrees, etc. Especially when it is pointed out that one's being morally virtuous in this sense does not *depend* in any way upon one's ability to *recognize goods as good*. One can be excellently concerned for things that are good—and so, other things equal, morally virtuous—even if one *neither* recognizes them as good *nor* even realizes that she is concerned for them. And as long as this is the case, then, per the argument in the last chapter, some will have moral intuitions that are trustworthy *because* they are grounded in emotion, and not despite this fact.

At one point, Street considers a response like this one, according to which our capacity to perceive objective moral values arose as a kind of byproduct of some *other* capacity that was selected for its evolutionary benefits. She calls this the “byproduct hypothesis”:

Many human capacities, after all, are like this: our ability to do astrophysics, for instance, was surely not itself directly selected for, but is instead the byproduct or outgrowth of other capacities which likely do have an explanation in terms of natural selection. Perhaps in some similar fashion our ability to grasp independent evaluative truths has emerged as a byproduct or outgrowth of some other capacity—call it *capacity C*.¹⁴

What I’m suggesting is that our ability to perceive objective moral values may very well have developed as a byproduct of our more basic capacities to have concerns and to respond to things emotionally.

Street objects to this sort of response by insisting that the same Darwinian dilemma arises once again: we must either *affirm* or *deny* a relationship between the evolution of capacity *C* and objective moral values. We should only *affirm* such a relation, though, if the resulting account of the evolution of *C* would be more acceptable scientifically than an account that does *not* affirm the relation. And Street doubts that this would be the case. On the other hand, if we *deny* a relation, Street thinks the coincidence worry arises again: if objective moral values had nothing whatsoever to do with the evolution of capacity *C*, then it would be *too coincidental to believe* that *C* could have given rise to a capacity to perceive objective moral values. However, as I hope is clear by now, this just isn’t the case. If moral virtue is roughly as I conceive it in the last chapter, then it is hardly a “complete fluke” that our capacity for emotion could give rise

¹⁴ Street 2006: 142.

to a capacity to perceive objective moral values, as Street claims it would have to be. As I said, all I need is for it to have been possible for us to develop concerns for things other than the satisfaction of our most basic evolutionary needs. And surely this has been the case, probably for as long as we've been able to have concerns at all.

Now, to be sure, Street would likely not let the discussion stop here. Presumably, the next step would be for her to ask: Assuming that many of us have concerns for things that are objectively morally good, what is the best explanation for *this* fact? And it seems the dilemma—or something like it—is up and running yet again. Unless moral values figure into our best explanation of how people have come to be sufficiently virtuous (as I am thinking of virtue), then it could only have been by a “complete fluke” that this has happened. But surely, there will be better explanations, by scientific standards, of the causal origins of our concerns than any involving reference to objective moral values.

I wonder, though, if this has now become a *false* dilemma. For it is just not clear to me that moral values would *have to* figure into our best explanation of why we have the concerns that we have—or else it is a “complete fluke”—as long as there's a plausible story according to which we *did in fact* develop concerns for things that are objectively morally good. What one needs are mechanisms whose existence and influence are affirmed by our best evolutionary sciences, but that can also figure into a plausible explanation of how it is that human beings could universally and systematically come to be concerned about things that are morally good. This way, the convergence of human concerns upon things that are good would not be a “complete fluke,” but nor would it be the case that the *goodness* of these things is necessary to explain how, or why, we came

to be concerned about them. And as far as I can tell, there *are* such mechanisms, well established by the relevant sciences. For instance, it is deeply ingrained in our evolved psychologies to be greatly concerned for our own well being, and for the well being of those close to us. And these are things that are among the goods for which a morally virtuous person is concerned. Furthermore, humans are well known to have a capacity to share, psychologically, in the suffering and enjoyment of others—a capacity known to most as *empathy*. This is something for which we have evolution to thank; and it is not difficult to imagine how this capacity could have greatly influenced our concerns for the avoidance of suffering and promotion of enjoyment (again, in ourselves and others). Now, to be sure, these are cases in which evolution is equipping us with concerns for things that are good, but not *because they are good*. But as I've explained, this is no matter to me. All that matters is that there is a plausible story to tell according to which human beings, evolved as we are, are concerned about things that are morally good (in a *de re* sense). Admittedly, much more than I've said here will be required for a full response to the sort of skepticism that Street encourages. But I do think I've said enough to give some indication of the avenue that I think one should take in order to do so.¹⁵

Both Greene and Street think that there is something about the evolution of emotions that tells against their being able to reliably track objective moral values.

¹⁵ On one reading of Street's argument, it is essentially just a restatement—perhaps with a few new evolutionary bells and whistles—of Gilbert Harman's famous argument against realist theories in ethics. Roughly, according to Harman, there is no phenomenon—emotional experiences, or otherwise—the best explanation of which includes reference to moral values (see Harman 1977 and 1986). In response, Nicholas Sturgeon rejected the supposed explanatory impotence of moral values, citing examples in which such values do seem to be making an explanatory difference that Harman apparently denies (see Sturgeon 1985 and 1986). As I say above, a full response to Street would require much more than I offer here, and likely take me too far away from my focus in this dissertation. But I mention the Harman/Sturgeon debate if for no other reason than to suggest it as a resource for those sentimentalist-realists like myself who find themselves troubled by Street-style skepticism.

Greene, however, appears only to beg certain very important questions. And Street's Darwinian dilemma, I think, turns out to be a *false* dilemma. According to her, the realist must either affirm or deny a relation between the forces of natural selection and objective moral values; and if he denies a relation (as I'm inclined to do), there is no way to avoid the conclusion that our moral intuitions are not trustworthy—since, if *any* of our moral intuitions are true, this could only be a matter of dumb luck, and not a reliable belief-forming process. But my accounts of emotion and moral judgment open up a third option. Regardless of how and why the emotions evolved, they *can* serve as a kind of perceptual basis for moral judgment, provided that they are shaped to a sufficient degree by moral virtue.

2. Nichols and the arationality of emotions

Shaun Nichols is another moral judgment skeptic. In his book *Sentimental Rules* (2004), Nichols develops a theory of moral judgment according to which moral judgment involves two dissociable mechanisms. The first of these mechanisms is an internally represented set of rules prohibiting certain behaviors—a “normative theory,” as he calls it. And the second is emotion. Psychopaths, he thinks, possess only the first of these. This is why they are often able to produce all of the “right” claims about what things are morally right and wrong, but nonetheless unable to recognize a difference between moral and conventional wrongs. In psychologically normal people, these internally represented sets of rules come to be “backed” by emotions like anger and disgust—hence, “sentimental rules”—and this is how we judge the moral wrongness of things. The details of his account are not important for my purposes here. What matters is only that Nichols, too, is a psychological sentimentalist.

After spending several chapters developing and defending his view, Nichols devotes the final chapter of his book to arguing that, *because it is a version of psychological sentimentalism*, his theory of moral judgment is incompatible with “commonsense [moral] objectivism.” As Nichols himself puts the argument:

Morality is not objective because:

1. Rational creatures who lack certain emotions would not make the moral judgments that we do.
2. There is no principled basis for maintaining that these certain emotions (on which our moral [intuitions] depend) are the *right* emotions. That is, there is no externally privileged basis for maintaining that all rational creatures *should* have the emotions.

In brief, moral judgment depends on certain emotions, and these emotions them-selves are rationally arbitrary, so moral judgment, as we know it, is not objective.¹⁶

Premise 1 is supposed to be an implication of psychological sentimentalism. And premise 2 is an assumption that Nichols makes about the nature of emotion, which I’ll address in a moment. I should also note that it is not entirely clear how Nichols is using the term ‘objective’ in this argument, since he applies the term to *morality* at the beginning, and then to *moral judgment* in the conclusion. This is unfortunate, since it is surely possible for *one* of these things to be objective, and not the other. Even if ordinary moral judgment in humans turns out not to be a process that can be counted upon to produce beliefs about objective moral values, it could of course still be the case that there *are* objective moral values. Maybe objective moral values are just inaccessible. Regardless, because of his focus upon the role of emotion in moral judgment (and also because of his rejection of *philosophical* versions of sentimentalism), I’ll henceforth assume that Nichols means to be making a claim just about the objectivity of *moral judgment*, and not about the objectivity of *morality* more generally. And that claim is this:

¹⁶ Nichols 2004: 185.

since emotion figures centrally in moral judgment, and the emotions are rationally arbitrary, moral judgment cannot be counted upon to produce beliefs about objective moral values.

In an attempt to explain how his argument is supposed to work, Nichols offers the following analogy. *We* believe that torturing puppies is wrong, but Martians who lack anything like human emotions might *not* believe this.

Furthermore, the argument continues, this need not be because the Martians are ignorant of any facts or because they are obtuse or too busy to do the calculations. Rather, the Martians might be as smart, well-informed, and have as much intellectual leisure as you please. The Martians might be exceptional (if solitary and back-stabbing) mathematicians and scientists. They fail to condemn puppy torture because they lack analogues of the human sentiments.¹⁷

So far, this is just supposed to be a case if Nichols' first premise. Here is where his second premise enters the analogy:

The next step in the broad argument is to maintain that there is no externally privileged basis for saying that we have the right emotions and the Martians are emotionally defective. That is, there is no independent basis for saying that our emotional repertoire is the right one to have.¹⁸

And from this it is supposed to follow that moral judgment is not objective.

I have two things to say in response to Nichols' argument. First, if the Aristotelian account of the relation between emotion and virtue described in the last chapter is on point, then Nichols is surely wrong to think that there is *no basis* upon which we might judge the rightness or wrongness of different emotional responses. Even if virtue is relative to *kind*—so that what makes for a virtuous *human* is something

¹⁷ Nichols 2004: 185.

¹⁸ Nichols 2004: 185-186.

entirely different from what makes for a virtuous *Martian*—it will still be false that “there is no principled basis for maintaining that these certain emotions ... are the *right* emotions.” That is, as long as we’re talking about *human* emotions. As I explained in chapter II, emotions are often susceptible to both rational and moral evaluation. Fear is the *right* response to a ferocious bear, and the *wrong* response to Yogi bear. Again, as I explained before, even neo-Jamesians like Prinz are careful to secure *some* sense in which emotions are susceptible to evaluation of this kind. So for Nichols to simply state, without any argument, that emotions are rationally arbitrary is surprising. And this is to say nothing of the *moral* evaluability of emotions, which is surely implied by our ordinary thought and talk about our own emotions and the emotions of others. If shame is not a morally appropriate response to serious wrongdoing, then apparently, we are all massively deceived about the nature of emotion.

Second, even if there *were* no basis for judging the rightness or wrongness of different emotional responses, it would not follow from this that the centrality of emotion in moral judgment is a reason to doubt that moral judgment is objective. Once again, it is useful to think of perception. Suppose I’m seated in a room with decent lighting, I’m wearing my glasses, my eyes are open, etc., and I have the perceptual experience of a coffee cup. On the basis of this experience, I come to believe that there is a coffee cup before me. My Martian friend, however, has perceptual faculties that are radically different from mine (he may also lack the relevant concepts), and consequentially has a different perceptual experience from the one that I have. Presumably, there is no privileged basis external to both of us upon which I could claim that *my* perceptual experience, *and not his*, is the right one to have. After all, his Martian faculties are

working exactly as Martian faculties should. But does it somehow follow from this that my perceptual belief does not result from a process that gives me access to objective features of reality (e.g., the fact that there is a coffee cup before me)? Of course not. So why think any differently about our moral intuitions and the emotions in which they are grounded?

Some may want to spin Nichols' argument into a more straightforwardly *epistemological* problem. One might think, for instance, that my inability to provide grounds for thinking that my perceptual experience, and not the experience had by my Martian friend, is the *right* one to have may nonetheless jeopardize the *justification* of my subsequent belief that there is a coffee cup before me. And if this is right, then the same sort of worry will apply to moral intuitions grounded in emotion. Unless we can access some privileged basis, external to both humans and Martians, upon which we might confirm the appropriateness of our emotional horror at puppy torture, then we are not justified in believing that it is wrong to torture puppies. At this point, my endorsement of process reliabilism (see §4 of chapter IV) is important. As long as our beliefs about the wrongness of puppy torture are reliably formed, then they are justified. The lack of an "externally privileged basis" in this case may certainly jeopardize our ability to justify our moral beliefs to our Martian friends. But *being justified in believing b* and *being able to justify b to others* are two importantly different things, especially as far as the reliabilist is concerned. I might be perfectly justified in believing *b*, even if I am wholly unable to explain my reasons for believing, or to articulate reasons for thinking the belief justified, or to persuade others of the belief.

3. Sinnott-Armstrong and the clouding effect of emotions

The final argument for moral judgment skepticism comes from Walter Sinnott-Armstrong. To set the argument up, Sinnott-Armstrong considers a variety of cases in which some inferential confirmation is needed in order for a *non*-moral belief to count as justified. And from these cases, he tries to draw out five general epistemological principles that he can then extend to moral intuitions. So for instance, he writes, “If you believe that George Washington never told a lie, and if this belief comes from a legend spread by Washington’s allies to gain power, then you are not justified in believing the legend, though it still might be true.”¹⁹ From this case, and others like it, Sinnott-Armstrong draws out the following principle: “confirmation is needed for a believer to be justified when the belief arises from an unreliable or disreputable source.”²⁰ Fair enough. His third principle, however, has to do with the impact of *emotion* upon the justification of beliefs. He writes,

When people get very angry, for example, they tend to overlook relevant facts. They often do not notice excuses or apologies by the person who made them angry. We should not generalize to all emotions, but we can still endorse some-thing like this:

Principle 3: confirmation is needed for a believer to be justified when the believer is emotional in a way that clouds judgment.²¹

So far, Sinnott-Armstrong hasn’t said anything that anyone should find objectionable. After all, there is no doubt that our emotions do *sometimes* cloud judgment. And he is even careful here to note that we should not generalize to *all* emotions. But later, when

¹⁹ Sinnott-Armstrong 2006: 345.

²⁰ Sinnott-Armstrong 2006: 345.

²¹ Sinnott-Armstrong 2006: 344.

he extends Principle 3 to apply to our ordinary moral intuitions, it becomes unclear whether he thinks that believers are ever *not* “emotional in a way that clouds judgment.”

Sinnott-Armstrong’s discussion of moral judgment focuses upon the following two questions. First, are moral intuitions primarily the products of reason or emotion? And second, which emotions, if any, cloud judgment? His answer to the first question is that moral intuitions are primarily the products of emotion, an endorsement of psychological sentimentalism. In fact, he briefly makes reference to many of the same studies that I discuss in chapter III. But what of the second question: Which emotions, if any, cloud moral judgment? Unfortunately, he never really clarifies what the difference is supposed to be between emotions that *do* cloud moral judgment and those that *do not*; and in fact, one gets the impression that he thinks emotions *always* cloud moral judgment. For instance, at one point, after referencing Greene’s fMRI study, in which subjects’ brains were scanned while they respond to the TROLLEY and FOOTBRIDGE scenarios, Sinnott-Armstrong writes,

It is not obvious what to make of these results. [...] Nonetheless, one natural speculation is this: When asked about pushing the fat man, subjects react, ‘That’s so horrible that I can’t even think about it.’ Emotions stop subjects from considering the many factors in these examples. If this interpretation is correct, then many pervasive and fundamental moral beliefs result from emotions that cloud moral judgment.²²

And a little later, he adds, “[S]uch brain studies seem to provide some evidence that many moral [intuitions] result from emotions that cloud judgment.”²³ Strictly speaking, though, neither Greene’s study nor any of the other studies that Sinnott-Armstrong cites makes any assumptions at all about the effects of emotion upon the *quality* of a person’s

²² Sinnott-Armstrong 2006: 351.

²³ Sinnott-Armstrong 2006: 352.

moral judgment—that is, none of these studies assumes that emotions are in the business of *clouding* moral judgment. All they purport to show is that many of our moral intuitions are causally grounded in emotion. So why would Sinnott-Armstrong think that these studies show that moral judgment is often *clouded* by emotion? Presumably, he draws this conclusion because he assumes that emotions *always* have a clouding effect—so that, to whatever extent moral judgment is based upon emotion, it is to that extent clouded.

Later still, Sinnott-Armstrong responds to the objection that, even if his five principles apply to *some* moral beliefs, they do not apply to *all* moral beliefs; and so there may still be moral beliefs that do not need inferential confirmation in order to be justified. In response, he writes,

As I admitted, some moral beliefs are not controversial. For example [referring to a variant of the FOOTBRIDGE case, described in chapter 3], almost everyone (except moral nihilists) agrees that it is morally wrong to push the fat man in front of the trolley just because you are angry with him for beating you in a game. Such cases also do not seem due to context, heuristics, overgeneralization, or framing effects. Still, such moral believers are [...] emotional (as Greene’s experiments suggested). So [Principle 3 does] seem to create a need for confirmation even in such clear cases.²⁴

Once again, Sinnott-Armstrong seems simply to assume that emotions are always in the business of clouding moral judgment. As long as “moral believers are emotional”—by which he seems to mean just that their beliefs are grounded in emotion—there is a “need for confirmation,” even if the belief itself is one to which almost everyone would agree.

Here, too, I have two things to say in response. First, while it surely must be conceded that emotions do sometimes cloud moral judgment, it also seems as if a *lack* of

²⁴ Sinnott-Armstrong 2006: 362

emotion can have similarly negative effects upon a person's judgment. As I discussed in chapter III, there are a number of psychological conditions characterized by, among other things, some sort of emotional deficit—i.e., a systematic absence or blunting of normal emotions. Psychopathy and autism are probably the two that receive the most attention in moral psychological literature. And as I explained in the earlier chapter, people with these conditions are well known for having serious difficulties with moral judgment. Psychopaths, for instance, are apparently unable to “see” a difference between moral and conventional wrongdoing, and most attribute this defect in their moral judgment to their characteristic emotional deficits. So while Sinnott-Armstrong is right to note, for instance, that our emotions sometimes keep us from considering features of our circumstances that may be relevant to moral judgment, there is also evidence for thinking that our emotions are sometimes *necessary* if those features are to ever be apparent to us in the first place.²⁵

And of course, the second thing to say in response to Sinnott-Armstrong is just that there are ways of conceiving the nature of emotion and its connection to morality—given, for instance, the relation that emotion bears to moral virtue—according to which it would be as mistaken to say that emotion clouds moral judgment as it would be to say that our perceptual experiences cloud our judgment of the perceptible features of things. It is hard to resist thinking that Sinnott-Armstrong (and perhaps other moral judgment skeptics) must be assuming some sort of *feeling theory* of emotion. Indeed, it would be easy to think that emotions could only be in the business of clouding moral judgment if

²⁵ People with autism, for instance, commonly struggle to recognize what is going on in the minds of other people, and many think that this struggle is an effect of “a failure to become emotionally engaged,” early in life, with the subjective states of others (Hobson 2005: 189). For more, see Hobson 2002.

one thought that emotions were nothing more than feelings of changes in the body. After all, when other such feelings are at the forefront of our minds, it seems likely that our judgment of things will consequently be less than ideal. Bouts of dizziness and heartburn, for instance, can be distracting in much the same way that Sinnott-Armstrong says emotions can be distracting.

However, for reasons that I lay out in chapter II, we should *not* think this way about the nature of emotion. There is more to emotions than just feelings. Emotions are complex states consisting of construals, concerns, and feelings, all bearing particular relations to each other. If I am right about this, and if the relation between emotion and moral virtue is roughly as I describe it in the last chapter, then the effect that virtue has in shaping our emotions is such as to make the virtuous person emotionally sensitive to objective moral values in a manner reminiscent of perception. And just as our fast and unreflective perceptual beliefs are (or can be) justified in virtue of their being grounded in perceptual experiences, our moral intuitions are (or can be) justified in virtue of, and not despite, their being grounded in emotional experiences. If my argument from chapter IV is successful, then, it turns out that Sinnott-Armstrong's assumption about the effects of emotion upon moral judgment is false—rather than only *clouding* moral judgment, it is *through* our emotions that we can become attuned to the moral values of things.

4. Sinnott-Armstrong on 'moral illusion'

In the same article, Sinnott-Armstrong offers another argument for doubting the trustworthiness of moral intuitions. This argument does not directly concern the nature or implications of the role of *emotion* in moral judgment, but I nonetheless think it will be

worth addressing here, for reasons that will become clear soon enough. Essentially, Sinnott-Armstrong argues that moral intuitions are often subject to what might be called “moral illusions,” and this is especially interesting since my case for the trustworthiness of moral intuitions in the last chapter involved an analogy to perception.

Here, too, his method is to begin with cases of *non*-moral beliefs that have been influenced in various ways by illusion, and then to work his way back to cases of moral beliefs. One of the non-moral examples is a case of visual illusion: objects look smaller when they are next to large objects, and larger when they are next to small objects. In such cases, then, Sinnott-Armstrong argues, “we are not justified in trusting our estimates until we check their sizes in other circumstances or by other methods.”²⁶

Another sort of illusion that he considers has to do with heuristics, i.e., quick-and-ready rules of thumb that unconsciously guide our thoughts in certain circumstances. For instance, if you ask people to guess how many times a 7-letter word with ‘n’ as the sixth letter appears in a particular book, they’ll typically give you a number *many times lower* than if you had asked them how many times a 7-letter word ending with ‘-ing’ appears in the book. But of course, any word of the latter type is also a word of the former type; so in reality, there must be at least as many words of the former type as there are of the latter type, if not more. Many attribute people’s responses here to an *availability heuristic*, according to which we judge the frequency of a thing based upon how easily examples of that thing come to mind: examples of 7-letter words ending in ‘-ing’ come to mind much more easily than examples of 7-letter words with ‘n’ as the sixth letter. But in this case, the heuristic has a kind of illusory effect—it makes 7-letter words with ‘n’ as the sixth

²⁶ Sinnott-Armstrong 2006: 344.

letter *appear* less common than they really are. As a result, Sinnott-Armstrong explains, people “do not seem adequately epistemically justified in trusting beliefs based on such heuristics.”²⁷

On the basis of cases like the two I’ve just mentioned, Sinnott-Armstrong suggests the following general principle: “confirmation is needed for a believer to be justified when the circumstances are conducive to illusion.”²⁸ And now we must ask: Are the circumstances of moral judgment ever conducive to illusion? Sinnott-Armstrong thinks they are.

One case of moral illusion comes from Peter Unger. As Sinnott-Armstrong recounts the case:

[Unger] found that the order in which options are presented affects beliefs about whether a given option is morally wrong. He also claims that people’s moral beliefs about a certain option depend on whether that option is presented as part of a pair or, instead, as part of a series that includes additional options intermediate between the original pair.²⁹

Another case of moral illusion comes from psychologists Kahneman and Tversky.³⁰ In one study, they divided subjects into two groups, asking members of each group to

²⁷ Sinnott-Armstrong 2006: 345.

²⁸ Sinnott-Armstrong 2006: 345.

²⁹ Sinnott-Armstrong 2006: 352 (citing Unger 1996: 88-94). As Sinnott-Armstrong explains in a footnote, Unger is here describing the results of an informal survey, but unfortunately, Unger does not describe the methods that he used for the survey. He also does not describe the *exact* results of the survey.

³⁰ I should say: It is not obvious that this should be counted as a case of *moral* illusion. Sinnott-Armstrong defends its inclusion in this discussion by arguing that, since subjects were not told how, if at all, the hypothetical programs would affect them personally, “their choices seem to result from beliefs about which program is morally right or wrong” (2006: 354).

choose between two hypothetical programs for fighting a particular disease that afflicts 600 people.³¹ Here is how the two programs were described to the first group:

If program A is adopted, 200 people will be saved. If program B is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved.

And here is how the two programs were described to subjects in the second group:

If program C is adopted, 400 people will die. If program D is adopted, there is a 1/3 probability that nobody will die, and a 2/3 probability that 600 will die.

Clearly, programs A and C are equivalent, and programs B and D are equivalent. But subjects' preferences did not reflect these equivalencies. When asked to choose between A and B, most subjects preferred A; and when asked to choose between C and D, most subjects preferred D. As Sinnott-Armstrong explains, despite the technical equivalence between A and C, for instance, the different terms of their descriptions create a kind of moral illusion: describing C in terms of how many people *will die*, as opposed to how many people *will be saved*, leads subjects to see it as (morally) worse than its equivalent. Illusions such as this one are sometimes classified as cases of *framing effects*, and Sinnott-Armstrong apparently believes they are quite common in moral judgment: "Framing effects distort moral beliefs in so many cases that moral believers need confirmation *for any particular moral belief*."³²

I have no intention to disagree with Sinnott-Armstrong's suggestion that moral judgment is sometimes subject to framing effects and other types of moral illusion. But I do wish to say two things in response. First, it seems odd to me that he characterizes the

³¹ Kahneman and Tversky 1979.

³² Sinnott-Armstrong 2006: 354; italics mine.

possibility of moral illusion as a *separate* reason for doubting the trustworthiness of moral intuitions—i.e., separate from their being grounded in emotion. Does he think that moral intuitions are grounded in emotion, or does he think that moral intuitions are grounded in something like an unconscious analysis of information and context that can sometimes be misled by things like framing effects? On the one hand, if he thinks that the evidence supports psychological sentimentalism, then shouldn't he want to interpret these latter studies in light of that fact? On the other hand, if he thinks most intuitions are subject to framing effects and other illusions, what does that have to do with emotion?

I raise these questions not because I think that Sinnott-Armstrong owes us answers to them, but rather as an opportunity to point out that *my* account of the nature of emotion and its role in moral judgment provides *me* with answers to them. Briefly: moral intuitions are grounded in emotion *and* they are sometimes subject to illusion, because what is happening in cases of moral illusion is that our *construals* of things are being manipulated, sometimes by factors that are not especially relevant to moral judgment. Subjects in the Kahneman and Tversky study are led to construe programs A and C differently; and this affects their subsequent (moral) beliefs about the two programs, because we respond differently emotionally to the thought of 400 people dying than we do to the thought of 200 people being saved. Similarly, in the case from Unger, it is likely that the order in which we are given options can affect our *construals of*—and subsequent moral *beliefs about*—the relevant options. So my account of emotion and its role in moral judgment provides me with a unifying explanation of what Sinnott-Armstrong treats as two disparate phenomena in moral psychology. If nothing else, I

take this to be further confirmation that I've got the anatomy of emotion and moral judgment basically right.

What can I say about the *trustworthiness* of moral intuitions, though, in light of the possibility of moral illusion? Essentially, I'll repeat what I said in chapter IV, in response to the worry that ordinary people are not nearly as virtuous as the MVP. Whereas the worry there was that our *concerns* are not always as they ought to be, here the worry seems to be that our *construals* of things are not always accurate. But the lessons are the same as they were before. First, as I said then, much more research needs to be done before anyone should join Sinnott-Armstrong in concluding that moral intuitions are not to be trusted. It simply does not follow from the cases that Sinnott-Armstrong cites that *all*, or even *most*, of our moral intuitions are influenced in this way. And even if it *can* be shown that moral judgment is often subject to illusion, skeptics will also have to show that the effects of these illusions are *great enough* to warrant skepticism. It will not be enough to simply point out that moral illusion causes our construals of things to be less than perfectly accurate. After all, even if my perceptions are never perfectly accurate, it may nonetheless be the case that my perceptual beliefs are reliably formed. With respect to ordinary perception, reliability does not entail perfect accuracy. Presumably, the same could be said of the construals that underlie our emotional responses to things.

Second, as before, there is a useful *practical* lesson to be taken from cases of moral illusion. Just as we need to take care with our concerns, being sensitive to cases in which our concerns may be otherwise than they ought to be, we need to take care with

our construals of things. This can be especially difficult, since, as I explained in chapter II, construals are sometimes unconscious, and sometimes out of our control. But nonetheless, we have a responsibility to do everything in our power to ensure that we see things as they are, morally. Take poverty, for instance. Most will agree that the suffering of those in conditions of poverty is a terrible thing. But it seems likely that those who have either experienced poverty for themselves, or have at least spent significant amounts of time with the poor, will construe this suffering *differently*—and arguably, *more accurately*—than those who have not. Similarly, those who spend time reflecting upon discrimination and its socio-psychological effects are likely to construe the circumstances of victims of discrimination more accurately than people who have never reflected upon such things. And so on and so forth. Developing habits like these may help to protect against the influence of moral illusions, by helping to ensure that we see things, morally, as they really are.

5. Conclusion

As I explained in the first chapter, moral intuitions are very important things, to philosophers and non-philosophers alike. And philosophers in particular have long debated whether or not these intuitions are truly deserving of the sort of trust that we apparently give to them, both in our ordinary lives and in moral philosophy. But until relatively recently, these debates involved very little attention to empirical research into the psychological underpinnings of moral judgment. This is unfortunate because, as Sinnott-Armstrong and others recognize, “it is hard to see how to determine whether certain moral intuitions are justified without any understanding of the processes that

produce those intuitions.”³³ Thankfully, interdisciplinary research into the nature of moral judgment and status of moral intuitions is now in full bloom, as moral philosophers, psychologists, neuroscientists, and others are jointly engaged in an effort to finally answer long-time philosophical questions in empirically adequate ways.

Unfortunately, the growing consensus seems to be that this research ought to *undermine* our trust in our own moral intuitions. This is unfortunate, for a number of reasons. For one thing, it would seem to imply that ethics ought to be done radically differently than philosophers have been doing it for the past two thousand years or so. For any and all references (direct or indirect) to moral intuitions—which are ubiquitous—would have to be completely disregarded, and avoided in the future. And for another thing, such sweeping skepticism of our moral intuitions would apparently have serious and disturbing implications upon the ways in which we conduct our everyday lives. As I explained in the first chapter, we not only regularly form moral beliefs about others and their actions in this fast and unreflective manner, but also, we organize our lives around, making decisions and plans on the basis of, these intuitive moral beliefs.

Underlying this growing consensus is the very popular assumption that moral intuitions are probably not trustworthy if they are grounded in emotion. Apparently, there is supposed to be something about *the nature of emotion* that somehow makes it unlikely, if not impossible, for beliefs grounded in emotion to ever be justified beliefs about objective moral values.

³³ Sinnott-Armstrong et al. 2010: 247.

In this dissertation, I've attempted to resist this trend by arguing that the popular assumption is false: moral intuitions *can* be trustworthy even if they are grounded in emotion. In fact, given the right conditions, moral intuitions can be trustworthy *because* they are grounded in emotion. As is sometimes said in other contexts, "the best defense is a good offense." So, in defense of the epistemic status of moral intuitions, I've gone on the offense, developing positive accounts not only of the nature of emotion, but also of the relation between emotion and moral virtue. If the nature of emotion is roughly as I argue in chapter II, and the relation between emotion and virtue is as I describe it in chapter IV, then it turns out that people's ordinary moral intuitions can be trustworthy *because* they are grounded in emotion, provided that these people are sufficiently morally virtuous.

Now, as I have been careful to recognize along the way, my overall account of moral judgment *raises* at least as many questions as it attempts to *answer*: Are there any empirical grounds for thinking that the nature of emotion is as I argue in chapter II? Might further research tell against psychological sentimentalism? How might we test for things like moral virtue, as I conceive them here? Is there any empirical basis for thinking that emotion relates to virtue in the way that I describe? And this is to say nothing of questions about any further *philosophical* implications that my views might have: What, if anything, might my view imply about the relation between sincere moral judgment and motivation? Does the fact that intuitions are grounded in emotion provide support for cognitivism, noncognitivism, or perhaps hybrid accounts of the function and meaning of moral utterances? Does my view imply that moral rules or principles play no

role at all in the production of moral intuitions? And if so, have I thereby committed myself to a form of moral particularism?

Unfortunately, all of these questions will have to be put off for a later time. My main concern in this dissertation has just been to construct an account of moral judgment that allows me to defend moral intuitions from the sorts of skeptical attacks that I describe here in chapter V. More work surely needs to be done, but I nonetheless think that the basic parts of a successful defense are all here.

REFERENCES

- ADAMS, R. M. (2006). *A Theory of Virtue*. Oxford: Oxford University Press.
- ALLADIN, A. (2008). *Cognitive Hypnotherapy: An Integrated Approach to the Treatment of Emotional Disorders*. West Sussex: Wiley and Sons.
- ARISTOTLE. *Nicomachean Ethics*. Trans. by W. D. Ross. In R. McKeon (ed.), *The Basic Works of Aristotle*. New York: Random House.
- ARPALY, N. (2003). *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford: Oxford University Press.
- BLAIR, R. J. (1993). "The Development of Morality." Ph.D. diss., University of London.
- (1995). "A Cognitive Developmental Approach to Morality: Investigating the Psychopath," *Cognition* 57: 1-29.
- (1996). "Brief Report: Morality in the Autistic Child," *Journal of Autism and Developmental Disorders* 26(5): 571-579.
- (2007). "The Amygdala and Ventromedial Prefrontal Cortex in Morality and Psychopathy," *Trends in Cognitive Sciences* 11: 387-392.
- BLAIR, R. J., ET AL. (1996). "Theory of Mind in the Psychopath," *Journal of Forensic Psychiatry* 7: 15-25.
- BLUM, L. A. (1994). *Moral Perception and Particularity*. Cambridge: Cambridge University Press.
- COLBY, A., KOHLBERG, L., GIBBS, J., AND LIEBERMAN, M. (1983). "A Longitudinal Study of Moral Judgment," *Monographs of the Society for Research in Child Development* 48.
- CUSHMAN, F., YOUNG, L., AND GREENE, J. D. (2010). "Multi-system Moral Psychology," in John M. Doris and the Moral Psychology Research Group (eds.), *The Moral Psychology Handbook*, Oxford: Oxford University Press, pp. 47-71.
- CUSHMAN, F., YOUNG, L., AND HAUSER, M. (2006). "The Role of Conscious Reasoning and Intuition in Moral Judgment: Testing Three Principles of Harm," *Psychological Science* 17(12): 1082-1089.
- DAMASIO, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: Gossett/Putnam.

- DARBY, B. W. AND JEFFERS, D. (1988). "The Effects of Defendant and Juror Attractiveness on Simulated Courtroom Trial Decisions," *Social Behavior and Personality* 16: 39-50.
- D'ARMS, J. (2005). "Two Arguments for Sentimentalism," *Philosophical Issues* 15: 1-21.
- D'ARMS, J. AND JACOBSON, D. (2006). "Sensibility Theory and Projectivism," in D. Copp (ed.), *Oxford Handbook of Ethical Theory*. New York: Oxford University Press.
- DION, K., BERSCHIED, E., AND WALSTER, E. (1972). "What is Beautiful is Good," *Journal of Personality and Social Psychology* 24: 285-290.
- DORIS, J. (2002). *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- DORIS, J. AND STICH, S. (2005). "As a Matter of Fact: Empirical Perspectives on Ethics," in Frank Jackson and Michael Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*. Oxford: Oxford University Press.
- DRETSKE, F. (1981). *Knowledge and the Flow of Information*. Cambridge: MIT Press.
- (1986). "Misrepresentation," in R. Bogdan (ed.), *Belief: Form, Content, and Function*. Oxford: Oxford University Press, pp. 17-36.
- DWYER, S. (2009). "Moral Dumbfounding and the Linguistic Analogy: Methodological Implications for the Study of Moral Judgment," *Mind & Language* 24(3): 274-296.
- DWYER, S., HUEBNER, B., AND HAUSER, M. (2010). "The Linguistic Analogy: Motivations, Results, and Speculations," *Topics in Cognitive Science* 2: 486-510.
- FOGAS, J. P. AND BOWER, G. H. (1987). "Mood Effects on Personal Perception Judgments," *Journal of Personality and Social Psychology* 51: 53-60.
- GIBBARD, A. (1990). *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- GOLDMAN, A. (2012). *Reliabilism and Contemporary Epistemology: Essays*. Oxford: Oxford University Press.
- GRANT, C. M., BOUCHER, J., RIGGS, K. J., AND GRAYSON, A. (2005). "Moral Understanding in Children with Autism," *Autism* 9(3): 317-331.
- GREENE, J. D. (2003). "From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology?" *Nature Reviews Neuroscience* 4: 847-850.

- (2008). “The Secret Joke of Kant’s Soul,” in Walter Sinnott-Armstrong (ed.), *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*. Cambridge, MA: MIT Press, pp. 35-79.
- GREENE, J. D., AND HAIDT, J. (2002). “How (and Where) Does Moral Judgment Work?” *Trends in Cognitive Sciences* 6: 517-523.
- GREENE, J. D., SOMMERVILLE, R. B., NYSTROM, L. E., DARLEY, J. M., AND COHEN, J. D. (2001). “An fMRI Investigation of Emotional Engagement in Moral Judgment,” *Science* 293: 2105-2108.
- GRIFFITHS, P. E. (2004). “Is Emotion a Natural Kind?” in R. Solomon (ed.), *Philosophers on Emotion*. Oxford: Oxford University Press, pp. 233-249.
- HAIDT, J. (2001). “The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment,” *Psychological Review* 108(4): 814-834.
- HAIDT, J., KOLLER, S. H., AND DIAS, M. G. (1993). “Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog?” *Journal of Personality and Social Psychology* 65(4): 613-628.
- HAIDT, J., BJORKLUND, F., AND MURPHY, S. (unpublished). “Moral Dumbfounding: When Intuition Finds No Reason.” University of Virginia.
- HARMAN, G. (1977). *The Nature of Morality: An Introduction to Ethics*. Oxford: Oxford University Press.
- (1986). “Moral Explanations of Natural Facts: Can Moral Claims Be Tested Against Moral Reality?” in N. Gillespie (ed.), *Spindel Conference: Moral Realism*. (*The Southern Journal of Philosophy*, suppl. vol. 24): 57-68.
- (1999). “Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error,” *Proceedings of the Aristotelian Society* 99: 315-331.
- HAUSER, M. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: HarperCollins.
- HOBSON, R. P. (2002). *The Cradle of Thought: Exploring the Origins of Thinking*. Oxford: Oxford University Press.
- (2005). “What Puts the Jointness in Joint Attention?” in N. Eilan, C. Hoerl, T. McCormack, and J. Roessler (eds.), *Joint Attention: Communication and Other Minds*. Oxford: Oxford University Press.

- HUME, D. (1983 [1751]). *An Enquiry Concerning the Principles of Morals*. J. B. Schneewind (ed.). Indianapolis: Hackett.
- (2000 [1739/1740]). *A Treatise of Human Nature*. D. F. Norton and M. J. Norton (eds.). Oxford: Oxford University Press.
- ISEN, A. M. AND LEVIN, P. F. (1972). "The Effect of Feeling Good on Helping: Cookies and Kindness," *Journal of Personality and Social Psychology* 21: 384-388.
- JAGGAR, A. (2008). "Love and Knowledge: Emotion in Feminist Epistemology," in Alison Jaggar (ed.), *Just Methods: An Interdisciplinary Feminist Reader*. Boulder: Paradigm Press, pp. 378-391.
- JAMES, W. (1950). *The Principles of Psychology, Vol. 2*. New York: Dover.
- KAHNEMAN, D. AND TVERSKY, A. (1979). "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47(2): 263-292.
- KAMTEKAR, R. (2004). "Situationism and Virtue Ethics on the Content of Our Character," *Ethics* 114: 458-491.
- KENNY, A. (1963). *Action, Emotion, and Will*. London: Routledge & Kegan Paul.
- KOHLBERG, L., LEVINE, C., AND HEWER, A. (1983). *Moral Stages: A Current Formulation and a Response to Critics*. In J. A. Meachem (Ed.), *Contributions to Human Development, Vol. 10*. Basel: Karger.
- LANGE, C. G. (1967 [1885]). "The Emotions: A Physiological Study," in C. Lange and W. James (eds.), *The Emotions*. New York: Hafner.
- LAZARUS, R. S. (1991). *Emotion and Adaptation*. New York: Oxford University Press.
- LERNER, J., GOLDBERG, J., AND TETLOCK, P. E. (1998). "Sober Second Thought: The Effects of Accountability, Anger, and Authoritarianism on Attributions of Responsibility," *Personality and Social Psychology Bulletin* 24: 563-574.
- LITTLE, M. O. (1995). "Seeing and Caring: The Role of Affect in Feminist Moral Epistemology," *Hypatia* 10: 117-137.
- MACKIE, J. L. (1977). *Ethics: Inventing Right and Wrong*. London: Penguin.
- MCDOWELL, J. (1998a). "Virtue and Reason," in J. McDowell (ed.), *Mind, Value, & Reality*. Cambridge, MA: Harvard University Press, pp. 50-73.
- (1998b). "Values and Secondary Qualities," in J. McDowell (ed.), *Mind, Value, & Reality*. Cambridge, MA: Harvard University Press, pp. 131-150.

- MEINONG, A. (1972). *On Emotional Presentation*. Trans. with Introduction by M. L. Schubert Kalsi. Evanston: Northwestern University Press.
- MICHAUD, S., AND AYNESWORTH, H. (1989). *Ted Bundy: Conversations with a Killer*. New York: New American Library.
- MOORS, A. (2010). "Theories of Emotion Causation: A Review," in J. De Houwer and D. Hermans (eds.), *Cognition & Emotion: Reviews of Current Research and Theories*. New York: Psychology Press.
- NICHOLS, SHAUN (2004). *Sentimental Rules: On the Natural Foundations of Moral Judgment*. New York: Oxford University Press.
- NISBETT, R. AND WILSON, T. (1977). "Telling More Than We Can Know: Verbal Reports on Mental Processes," *Psychological Review* 84(3): 231-259.
- OAKLEY, J. (1992). *Morality and the Emotions*. London: Routledge.
- PIAGET, J. (1997). *The Moral Judgment of the Child*. New York: Free Press.
- PRINZ, J. (2000). "The Duality of Content," *Philosophical Studies* 100: 1-34.
- (2002). *Furnishing the Mind: Concepts and their Perceptual Basis*. Cambridge, MA: MIT Press.
- (2004). *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press.
- (2006). "The Emotional Basis of Moral Judgments," *Philosophical Explorations* 9(1): 29-43.
- (2007). *The Emotional Construction of Morals*. Oxford: Oxford University Press.
- ROBERTS, R. C. (1988). "What An Emotion Is: A Sketch," *Philosophical Review* 97(2): 183-209.
- (1989). "Aristotle on Virtues and Emotions," *Philosophical Studies* 56(3): 293-306
- (2003). *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press.
- (unpublished). "Emotions, Perceptions, and Moral Judgments."

- RORTY, A. (1980). "Explaining Emotions," in A. Rorty (ed.), *Explaining Emotions*. Berkeley: University of California Press, pp. 103-126.
- SCHNALL, S., HAIDT, J., CLORE, G. L., AND JORDAN, A. (2008). "Disgust as Embodied Moral Judgment," *Personality and Social Psychology Bulletin* 34: 1096-1109.
- SHAFFER-LANDAU, R. (2003). *Moral Realism: A Defence*. Oxford: Clarendon Press.
- (2004). *Whatever Happened to Good and Evil?* Oxford: Oxford University Press.
- SINGER, P. (1972). "Famine, Affluence and Morality," *Philosophy and Public Affairs* 1: 229-243.
- (1981). *The Expanding Circle: Ethics and Sociobiology*. Oxford: Oxford University Press.
- (2007). "Should We Trust Our Moral Intuitions?" *Project Syndicate: A World of Ideas*, March 14, 2007. Accessed February 17, 2013. <http://www.project-syndicate.org/commentary/should-we-trust-our-moral-intuitions>.
- SINNOTT-ARMSTRONG, W. (2006). "Moral Intuitionism Meets Empirical Psychology," in T. Horgan and M. Timmons (eds.), *Metaethics After Moore*. Oxford: Clarendon Press, pp. 339-366.
- SINNOTT-ARMSTRONG, W., YOUNG, L., AND CUSHMAN, F. (2010). "Moral Intuitions," in John M. Doris and the Moral Psychology Research Group (eds.), *The Moral Psychology Handbook*, Oxford: Oxford University Press, pp. 246-272.
- SMETANA, J. (1993). "Understanding of Social Rules," in M. Bennett (ed.), *The Development of Social Cognition: The Child as Psychologist*. New York: Guilford Press.
- SMETANA, J. AND BRAEGES, J. (1990). "The Development of Toddlers' Moral and Conventional Judgments," *Merrill-Palmer Quarterly* 36: 329-346.
- SOLOMON, R. (1980). "Emotions and Choice," in Amélie Rorty (ed.), *Explaining Emotions*. Berkeley: University of California Press, pp. 251-282.
- SOMMERS, T. (2009). *A Very Bad Wizard: Morality Behind the Curtain: Nine Conversations*. McSweeney's Press.
- STRACK, F., MARTIN, L. L. AND STEPPER, S. (1988). "Inhibiting and Facilitating Conditions of Facial Expressions: A Nonobtrusive Test of the Facial Feedback Hypothesis," *Journal of Personality and Social Psychology* 54: 768-777.

- STRATTON-LAKE, P. (2002). "Introduction," in P. Stratton-Lake (ed.), *Ethical Intuitionism: Re-evaluations*. Oxford: Clarendon Press, pp. 1-28.
- STREET, S. (2006). "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127(1): 109-166.
- STURGEON, N. (1985). "Moral Explanations," in D. Copp and D. Zimmerman (eds.), *Morality, Reason and Truth*. Totowa NJ: Rowman & Allanheld, pp. 49-78.
- (1986). "Harman on Moral Explanations of Natural Facts," in N. Gillespie (ed.), *Spindel Conference: Moral Realism*. (*The Southern Journal of Philosophy*, suppl. vol. 24): 69-78.
- TAKEDA, T., KASAI, K., AND KATO, N. (2007). "Moral Judgment in High-Functioning Pervasive Developmental Disorders," *Psychiatry and Clinical Neurosciences* 61: 407-414.
- UNGER, P. (1996). *Living High and Letting Die*. New York: Oxford University Press.
- VALDESOLO, P., AND DESTENO, D. (2006). "Manipulations of Emotional Context Shape Moral Judgment," *Psychological Science* 17: 476-477.
- WARNOCK, G. J. (1967). *Contemporary Moral Philosophy*. London: Palgrave Macmillan.
- WEDGWOOD, R. (2007). *The Nature of Normativity*. Oxford: Clarendon Press.
- WHEATLEY, T. AND HAIDT, J. (2005). "Hypnotically Induced Disgust Makes Moral Judgments More Severe," *Psychological Science* 16: 780-784.
- WILSON, T. AND NISBETT, R. (1978). "The Accuracy of Verbal Reports About the Effects of Stimuli on Evaluations and Behavior," *Social Psychology* 41(2): 118-131.
- WITTGENSTEIN, L. (1953). *Philosophical Investigations*. Trans. by G. E. M. Anscombe. Oxford: Blackwell.
- YOUNG, L., BECHARA, A., TRANEL, D., DAMASIO, H., HAUSER, M., AND DAMASIO, A. (2010). "Damage to Prefrontal Cortex Impairs Judgment of Harmful Intent," *Neuron* 65: 845-851.
- ZAGZEBSKI, L. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.
- ZAJONC, R. B., MURPHY, S. T., AND MCINTOSH, D. N. (1989). "Feeling and Facial Efference: Implications of the Vascular Theory of Emotion," *Psychological Review* 96: 395-416.